Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **400,** 851–865 (2009) doi:10.1111/j.1365-2966.2009.15490.x

# Covariance of cross-correlations: towards efficient measures for large-scale structure

Robert E. Smith★

*Institute for Theoretical Physics, University of Zurich, Zurich CH 8037, Switzerland*

Accepted 2009 August 4. Received 2009 July 17; in original form 2009 June 13

**ABSTRACT**

We study the covariance of the cross-power spectrum of different tracers for the large-scale structure. We develop the counts-in-cells framework for the multitracer approach, and use this to derive expressions for the full non-Gaussian covariance matrix. We show that for the usual autopower statistic, besides the off-diagonal covariance generated through gravitational mode-coupling, the discreteness of the tracers and their associated sampling distribution *can* generate strong off-diagonal covariance, and that this becomes the dominant source of covariance as spatial frequencies become larger than the fundamental mode of the survey volume. On comparison with the derived expressions for the cross-power covariance, we show that the off-diagonal terms can be suppressed, if one cross-correlates a high tracer-density sample with a low one. Taking the *effective* estimator efficiency to be proportional to the signal-to-noise ratio (S/N), we show that, to probe clustering as a function of physical properties of the sample, i.e. cluster mass or galaxy luminosity, the cross-power approach can outperform the autopower one by factors of a few. We confront the theory with measurements of the mass–mass, halo–mass and halo–halo power spectra from a large ensemble of $N$-body simulations. We show that there is a significant S/N advantage to be gained from using the cross-power approach when studying the bias of rare haloes. The analysis is repeated in configuration space and again S/N improvement is found. We estimate the covariance matrix for these samples, and find strong off-diagonal contributions. The covariance depends on halo mass, with higher mass samples having stronger covariance. In agreement with theory, we show that the covariance is suppressed for the cross-power. This work points the way towards improved estimators for studying the clustering of tracers as a function of their physical properties.

**Key words:** cosmology: theory – large-scale structure of Universe.

## 1 INTRODUCTION

The power spectrum of matter fluctuations is of prime concern in cosmology, since it contains detailed information about the underlying world model and provides a method for probing the initial conditions of the Universe. Moreover, if the statistical properties of the initial fluctuations form a Gaussian random field, as is the case for most inflationary models, then the power spectrum provides a complete description for the spatial statistics of the density field. Consequently, over the last few decades a large fraction of observational and theoretical effort has been invested in estimating the power spectrum of galaxies from large redshift surveys and also to devising methods for extracting cosmological information from the signal (Feldman, Kaiser & Peacock 1994; Peacock & Dodds 1994; Percival et al. 2001; Tegmark et al. 2004b; Cole et al. 2005; Tegmark et al. 2006; Percival et al. 2007).

In order to obtain robust cosmological constraints from such data sets, one, however, requires additional knowledge about the signal covariance matrix – or the correlation function of power fluctuations. Unlike the power spectrum, which is the Fourier transform of the two-point correlation function, the covariance has had relatively little attention. This mainly stems from the fact that in order to estimate this quantity from a galaxy survey, or to compute it theoretically, one is required to investigate the four-point function of Fourier modes, more commonly the trispectrum of galaxies, and this is a substantially more complex quantity.

The first study of power spectrum covariance, in the modern context, was performed by Feldman et al. (1994), who showed, under the assumption of Gaussianity, that the matrix was diagonal and that the variance per band power was proportional to the square of the power in the band (see also Stirling & Peacock 1996). This result gave impetus to those advocating the use of power spectra

★E-mail: res@physik.unizh.ch

for large-scale structure work, over the simpler two-point correlation function, $\xi$, since under these same assumptions $\xi$ possesses correlated errors (Bernstein 1994).

Later, both Scoccimarro, Zaldarriaga & Hui (1999) and Meiksin & White (1999) independently showed that the real situation was much more complicated than the Gaussian calculation would lead one to believe. They recognized that non-linear gravitational instability would cause different Fourier modes to become coupled together, thus breaking the Gaussianity. In Scoccimarro et al. (1999), this mode-coupling behaviour was demonstrated by using higher order perturbation theory to calculate the trispectrum and by an analysis of results from an ensemble of $N$-body simulations. One direct consequence of this effect was that the fractional errors on the dark matter power spectrum were shown to reach an almost constant plateau on intermediate to small scales, regardless of the additional number of Fourier modes (see also Scoccimarro & Sheth 2002; Hamilton, Rimes & Scoccimarro 2006; Rimes & Hamilton 2006; Takahashi et al. 2009). They also showed that off-diagonal covariance on small scales was generated, but their results on large scales appeared inconclusive, owing to small volumes and hence increased sample variance.

Meiksin & White (1999) reached similar conclusions. They also extended the theoretical analysis to include the covariance in the power spectrum, arising from the finite sampling of the density field, referred to as Poisson sampling variance. It is well known that this is of importance for correctly determining the diagonal errors of the covariance matrix for rare tracers of the density field, such as bright galaxies and clusters. Whilst the covariance matrix of the dark matter power spectrum has been studied in some detail, that of haloes and galaxies has not received nearly the same level of attention – at least not beyond the assumption of linear density evolution and linear biasing. Notable contributions are: Cooray & Hu (2001); Scoccimarro & Sheth (2002); Sefusatti et al. (2006); Angulo et al. (2008a). However, as we will show for the first time in this work, the discreteness terms that were neglected by Meiksin & White (1999, since they were mainly studying the dark matter clustering), inevitably, become the dominant source of off-diagonal error for discrete tracers of the mass distribution.

Recently, cross-correlation techniques have become an ever more important tool for extracting information from large-scale structure data. For instance, in a recent theoretical study, Smith, Scoccimarro & Sheth (2007) demonstrated, using $N$-body simulations, that the cross-power spectrum between dark matter and haloes had several advantages over the simpler autopower spectrum method. In particular, a reduced shot-noise correction and noise properties. This cross-correlation approach has been further exploited to elucidate the environmental dependence of halo bias (Jing, Suto & Mo 2007; Angulo, Baugh & Lacey 2008b) and recently as a means for probing the large-scale scale dependence of bias in models of primordial Non-Gaussianity (Dalal et al. 2008; Pillepich, Porciani & Hahn 2008; Desjacques, Seljak & Iliev 2009; Grossi et al. 2009). Also, the cross-correlation approach has recently been applied to real survey data: Padmanabhan et al. (2008) study the intrinsic clustering properties of quasars in the Sloan Digital Sky Survey photometric redshift catalogue, through cross-correlating them with the more abundant Luminous Red Galaxy (LRG) sample; Wake et al. (2008) apply similar analysis to the clustering of radio galaxies at $z \sim 0.5$ from the 2dF-SDSS LRG QSO (2SLAQ) survey. It is therefore of great use to have an explicit calculation for the covariance of cross-correlations for use in likelihood analysis. Moreover, the covariance matrix is an important ingredient for any Fisher matrix parameter forecast, and hence an essential tool for optimal survey design (Tegmark 1997).

The paper is organized as follows. In Section 2, we develop the standard counts-in-cells framework to calculate the cross-power spectrum of two different tracers of the large-scale structure. In the analysis, we pay close attention to the assumed sampling distribution: besides the usual Poisson model, we also consider the toy-model scenario where one tracer is simply a subsample of the other and results are presented for both the cases. This is instructive, since it is likely that not all galaxies are equally good tracers of the mass – in particular those hosted in the same halo. Then in Section 3, we derive an expression for the covariance of the cross-power spectrum, including all non-Gaussian and finite sampling contributions to the error. Limiting cases are considered and expressions are also given for band-power averages. We evaluate the expected covariance signal for several different tracers of the mass. In Section 4, we compare the efficiency of the cross-power approach with that of the simpler autopower approach. In Section 5, the analogous expressions are derived for the cross-correlation function. In Section 6, we make a direct comparison of the theoretical predictions with estimates measured from the ZHORIZON simulations, a large ensemble of dark matter $N$-body simulations with total volume $\sim 100\,h^{-3}\,\mathrm{Gpc}^3$. Finally, in Section 7 we summarize our results and conclude.

When this paper was in the refereeing stage, two related works appeared: Hütsi & Lahav (2008) used the Halo Model and a Fisher matrix approach to investigate improvements in parameter constraints to be gained from cross-correlation analysis and White, Song-S. & Percival (2008) also looked at information to be gained from cross-correlating multiple populations in redshift space, developing further the study of Seljak (2009) and McDonald & Seljak (2008).

## 2 COUNTS-IN-CELLS FRAMEWORK FOR MULTIPLE TRACERS

### 2.1 Statistics of a single tracer population

Consider a single population of $N$ discrete objects in some large volume $V_\mu$ that trace the large-scale structure of the Universe in some way. Following Peebles (1980), we will assume that these tracers are Poisson sampled from some underlying smooth density field, and that the statistics of this underlying field are well described by a Gaussian Random Field. Hence, on partitioning space into a set of infinitesimal volume elements $\delta V$, the probability of finding $N_i$ galaxies in an element at position vector $\boldsymbol{r}_i$ is given by

$$P(N_i|\lambda = n(\boldsymbol{r}_i)\delta V) = \frac{\exp(-\lambda)\lambda^{N_i}}{N_i!} \approx \begin{cases} n(\boldsymbol{r}_i)\delta V & (N_i = 1) \\ 1 - n(\boldsymbol{r}_i)\delta V & (N_i = 0)\,, \\ 0 & (N_i > 1) \end{cases} \quad (1)$$

where $n(\boldsymbol{r})$ is the continuous number density function for tracers in the volume, which, in the local model for galaxy bias (Fry & Gaztanaga 1993; Coles 1993), is directly related to the underlying distribution of fluctuations in the cold dark matter (CDM) and for the linearized relation this is simply: $n(\boldsymbol{r}) = \bar{n}[1 + b\delta(\boldsymbol{r})]$, where $\delta(\boldsymbol{r}) = [\rho(\boldsymbol{r}) - \bar{\rho}]/\bar{\rho}$ is the fractional overdensity in the dark matter relative to the mean density $\bar{\rho}$. The probabilities of finding $N_i \geq 2$ are higher powers of the infinitesimal quantity $\delta V$ and so are negligible.

Owing to the fact that the occupation probability $P(N|\lambda)$ is non-zero only when $N = 0$ or 1, all of the one-point moments can be easily derived and for ($m \geq 1$)

$$\left\langle N_i^m \right\rangle_{p,s} = \left\langle \sum_{N=0}^{\infty} P(N|\lambda_i) N^m \right\rangle_s = \langle n(\boldsymbol{r}_i)\delta V \rangle_s = \bar{n}\delta V = \langle N_i \rangle_{p,s}\,; \tag{2}$$

and the central moments of the distribution are ($m > 1$)

$$\langle (N_i - \langle N_i \rangle)^m \rangle_{p,s} = \langle n(\boldsymbol{r}_i)\delta V \rangle_s = \bar{n}\delta V, \tag{3}$$

where in the above we used the notation $\langle \ldots \rangle_{p,s}$ to denote an averaging over all possible samplings of the points $p$ and all points in space $s$ (for brevity we will simply write $\langle \ldots \rangle$).

The two-point moments may also be derived. Consider the joint probability of finding objects in two disjoint volume elements $\delta V_i$ and $\delta V_j$ separated by a vector $\boldsymbol{r}_{ij} = \boldsymbol{r}_i - \boldsymbol{r}_j$; in the Poisson sampling model this is given simply by the product of the independent probabilities ($i \neq j$):

$$P(N_i, N_j) = P(N_i)P(N_j) \tag{4}$$

$$= n(\boldsymbol{r}_i)n(\boldsymbol{r}_j)\delta V_i \delta V_j. \tag{5}$$

On averaging, the two-point moments may be written:

$$\left\langle N_i^k N_j^m \right\rangle = \bar{n}^2 \delta V_i \delta V_j [1 + \xi(\boldsymbol{r}_i, \boldsymbol{r}_j)], \tag{6}$$

where $\bar{n} \equiv \langle n(\boldsymbol{r}) \rangle = \sum_i N_i / V_\mu = N/V_\mu$ is the mean number density of tracers and $\xi(\boldsymbol{r}_i, \boldsymbol{r}_j)$ is the two-point autocorrelation function. Hence, correlations are introduced into the sample, if and only if the points in the underlying continuous field are correlated.

### 2.2 The autopower spectrum

We define the Fourier relations for the density field as

$$\delta(\boldsymbol{r}) = \frac{V_\mu}{(2\pi)^3} \int d^3k\, \delta(\boldsymbol{k}) \exp(-\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}), \tag{7}$$

$$\delta(\boldsymbol{k}) = \frac{1}{V_\mu} \int d^3r\, \delta(\boldsymbol{r}) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}). \tag{8}$$

The density field of the discrete counts in cells is written:

$$\delta^d(\boldsymbol{r}) = \frac{1}{\bar{n}} \sum_i (N_i - \langle N_i \rangle)\, \delta^D(\boldsymbol{r} - \boldsymbol{r}_i), \tag{9}$$

which on insertion into our definition of the Fourier transform leads to the discrete sum

$$\delta^d(\boldsymbol{k}) = \frac{1}{N} \sum_i (N_i - \langle N_i \rangle) \exp(\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}_\mathrm{i}). \tag{10}$$

We may now compute the power spectrum of the discrete set of tracers,

$$\left\langle \delta^d(\boldsymbol{k}_1)\delta^d(\boldsymbol{k}_2) \right\rangle = \frac{1}{N^2} \sum_{i,j} \langle [N_i - \langle N_i \rangle][N_j - \langle N_j \rangle] \rangle \times \mathrm{e}^{\mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{r}_\mathrm{i} + \mathrm{i}\boldsymbol{k}_2 \cdot \boldsymbol{r}_\mathrm{j}} \tag{11}$$

$$= \frac{1}{{V_\mu}^2} \sum_{i \neq j} \delta V_i \delta V_j \xi(\boldsymbol{r}_i, \boldsymbol{r}_j) e^{\mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{r}_\mathrm{i} + \mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{r}_\mathrm{j}} + \frac{1}{N V_\mu} \sum_{i=j} \delta V_i e^{\mathrm{i}(\boldsymbol{k}_1 + \boldsymbol{k}_2) \cdot \boldsymbol{r}_\mathrm{i}}. \tag{12}$$

The sums over cells can be transformed into volume integrals, and the double volume integral over the correlation function in the first term can be simplified by recalling that through statistical homogeneity $\xi(\boldsymbol{r}_i, \boldsymbol{r}_j) = \xi(\boldsymbol{r}_i - \boldsymbol{r}_j, 0)$. We may then make use of the orthogonality of the Fourier basis functions to evaluate sums of the type

$$\sum_i \delta V_i e^{\mathrm{i}(\boldsymbol{k}_1 + \boldsymbol{k}_2) \cdot \boldsymbol{r}_\mathrm{i}} = V_\mu \delta^K_{\boldsymbol{k}_1, -\boldsymbol{k}_2}. \tag{13}$$

Hence, after performing these steps and introducing our definition of the power spectrum as

$$P(\boldsymbol{k}_1)\delta^K_{\boldsymbol{k}_1, -\boldsymbol{k}_2} \equiv V_\mu \langle \delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2) \rangle, \tag{14}$$

we recover the standard result for the power spectrum of discrete tracers (Peebles 1980):

$$P^d(k) = P^c(k) + \frac{1}{\bar{n}}, \tag{15}$$

where $P^c$ is the power spectrum of the underlying continuous field of tracers. The constant term on the right-hand side of the equation is more commonly referred to as the 'shot-noise correction' term, and is the additional variance introduced through discreteness.

### 2.3 Statistics of two tracer populations

We will now extend the above formalism to the problem of two different tracer populations, which we will denote as $A$ and $B$. Let the total number of objects in samples $A$ and $B$ be $N_A$ and $N_B$, and the numbers of each type of object in the $i$th cell be $N_{A,i} \equiv N_A(\boldsymbol{r}_i)$ and $N_{B,i} \equiv N_B(\boldsymbol{r}_i)$, respectively. Likewise, the mean number densities are $\bar{n}_A$ and $\bar{n}_B$. We now consider two cases for the sampling distributions, these are as follows.

(i) *Non-overlapping tracers*. $A$ and $B$ are both independent Poisson samples of the underlying continuous density field. In this case, the joint probability distribution for obtaining objects of types $A$ and $B$ in a single cell is

$$P(N_{A,i}, N_{B,i}) = P(N_{A,i}|\lambda_A)P(N_{B,i}|\lambda_B),$$

$$\approx \begin{cases} 1 - [n_A(\boldsymbol{r}) + n_B(\boldsymbol{r})]\,\delta V & (N_A = 0, N_B = 0) \\ n_A(\boldsymbol{r})\delta V & (N_A = 1, N_B = 0) \\ n_B(\boldsymbol{r})\delta V & (N_A = 0, N_B = 1) \\ 0 & (N_A \geq 1, N_B \geq 1). \end{cases} \tag{16}$$

The one-point cross-moments are then calculable ($m \geq 1, k \geq 1$),

$$\left\langle N_{A,i}^m N_{B,i}^k \right\rangle = 0; \tag{17}$$

and so also the central moments:

$$\left\langle (N_{A,i} - \langle N_{A,i} \rangle)^m (N_{B,i} - \langle N_{B,i} \rangle)^k \right\rangle = 0. \tag{18}$$

As in equation (5), the two-point cross-moments may also be derived and these are ($i \neq j$)

$$\left\langle N_{A,i}^m N_{B,j}^k \right\rangle = \bar{n}_A \bar{n}_B \delta V_i \delta V_j [1 + \xi^{AB}(\boldsymbol{r}_i, \boldsymbol{r}_j)], \tag{19}$$

where $\xi^{AB}$ is the two-point cross-correlation function of the tracers $A$ and $B$.

(ii) *Overlapping tracers*. $A$ is a Poisson sample of the underlying continuous density field and $B$ is a subsample of $A$. This time the joint probability distribution for obtaining objects of types $A$ and $B$ is written:

$$P(N_{A,i}, N_{B,i}) = P(N_{A,i}|\lambda_A)P(N_{B,i}|N_{A,i}). \tag{20}$$

The conditional probability $P(N_{B,i}\,|N_{A,i})$ is the key object of interest here, and as a simple illustrative example we will take this as

$$P(N_{B,i}|N_{A,i}) = \begin{cases} 1 & (N_B = 0|N_A = 0) \\ a & (N_B = 1|N_A = 1) \\ 1-a & (N_B = 0|N_A = 1) \\ 0 & (N_B > 1|N_A \geq 1), \end{cases} \tag{21}$$

where we will fix $a \equiv N_B/N_A$. Again, the one-point cross-moments are also calculable for this sampling model,

$$\left\langle N_{A,i}^m N_{B,i}^k \right\rangle = a\,\bar{n}_A \delta V = \bar{n}_B \delta V; \tag{22}$$

and so also the central moments:

$$\left\langle (N_{A,i} - \langle N_{A,i}\rangle)^m (N_{B,i} - \langle N_{B,i}\rangle)^k \right\rangle = \bar{n}_B \delta V. \tag{23}$$

Similarly, the two-point cross-moments are also calculable,

$$\left\langle N_{A,i}^m N_{B,j}^k \right\rangle = \bar{n}_A \bar{n}_B \delta V_i \delta V_j \left[1 + \xi^{AB}(\boldsymbol{r}_i, \boldsymbol{r}_j)\right]. \tag{24}$$

### 2.4 The cross-power spectrum

We may also compute the cross-power spectrum of tracers $A$ and $B$, and for both the non-overlapping (case i) and overlapping (case ii) sampling distributions. The Fourier modes for tracers $A$ and $B$ can be written:

$$\delta_T^d(k) = \frac{1}{N_{\rm T}} \sum_i \left[N_{\rm T}(\boldsymbol{r}_i) - \left\langle N_{{\rm T},i}\right\rangle\right] \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}_{\rm i}), \tag{25}$$

where $\mathrm{T} = \{A, B\}$ denotes the tracer type. As for the autospectrum, we will define the cross-power spectrum as

$$P_{AB}(\boldsymbol{k}_1)\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_2} \equiv V_\mu \left\langle \delta_A(\boldsymbol{k}_1)\delta_B(\boldsymbol{k}_2)\right\rangle. \tag{26}$$

Following now the steps in Section 2.2, but this time using the statistics for the counts in cells as given in the previous section, we find that the cross-power of discrete tracers $A$ and $B$ obeys the relation

$$P^d_{AB}(k) = P^c_{AB}(k) + \left\{\frac{1}{\bar{n}_A}\right\}, \tag{27}$$

where $P^c_{AB}$ is the cross-power spectrum of the underlying continuous fields. This expression is almost identical to the result for the autospectrum (equation 15); however, we emphasize an important difference – the constant term is enclosed by curly brackets. *In this paper* $\{\ldots\}$ *shall have the special meaning that this term only appears when there is an overlap between samples A and B, as in sampling case (ii), otherwise this term is exactly zero (see Peebles 1980).* We note that this notation shall be exploited throughout the rest of the paper, to represent the results from both sampling distributions with a single equation. More intuitively, the appearance of the constant term in the cross-power spectrum warns us that, if the two samples are not truly independent then we should expect some finite sampling correction.

## 3 COVARIANCE OF THE CROSS-POWER SPECTRUM

We now turn to the calculation of the full non-Gaussian covariance of the cross-power spectrum for discrete tracers $A$ and $B$. Note that when considering sampling case (ii), and in the limit that $N_A = N_B$, we will recover the standard covariance relations for the autopower spectrum (Meiksin & White 1999; Scoccimarro et al. 1999).

### 3.1 Definition of the covariance

To begin with, we define the covariance, per mode, of the cross-power spectrum for discrete tracers $A$ and $B$ as

$$\begin{aligned} C^d_{AB} &\equiv \mathrm{Cov}\left[P^d_{AB}(\boldsymbol{k}_1),\, P^d_{AB}(\boldsymbol{k}_2)\right] \\ &= \left\langle P^d_{AB}(\boldsymbol{k}_1)P^d_{AB}(\boldsymbol{k}_2)\right\rangle - \left\langle P^d_{AB}(\boldsymbol{k}_1)\right\rangle\left\langle P^d_{AB}(\boldsymbol{k}_2)\right\rangle \end{aligned} \tag{28}$$

On inserting the definition for the cross-power spectrum, $P_{AB} \equiv V_\mu \langle \delta_A(\boldsymbol{k}_1)\delta_B(-\boldsymbol{k}_1)\rangle$, and making use of equation (27) in the second term on the right-hand side, we obtain

$$C^d_{AB} = {V_\mu}^2 \left\langle \delta^d_A(\boldsymbol{k}_1)\delta^d_B(-\boldsymbol{k}_1)\delta^d_A(\boldsymbol{k}_2)\delta^d_B(-\boldsymbol{k}_2)\right\rangle \tag{29}$$

$$- \left[P^c_{AB}(\boldsymbol{k}_1) + \left\{\frac{1}{\bar{n}_A}\right\}\right]\left[P^c_{AB}(\boldsymbol{k}_2) + \left\{\frac{1}{\bar{n}_A}\right\}\right]. \tag{30}$$

Thus, we see that in order to compute the covariance of the cross-power spectrum it is also necessary to evaluate the four-point function of Fourier modes, or more commonly the trispectrum.

### 3.2 Evaluating the discrete cross-trispectrum

Using the counts-in-cells approach, the four-point cross-correlation function of Fourier modes can be written explicitly as

$$\begin{aligned} \left\langle \delta^d_A(\boldsymbol{k}_1)\delta^d_A(\boldsymbol{k}_2)\delta^d_B(\boldsymbol{k}_3)\delta^d_B(\boldsymbol{k}_4)\right\rangle &= \frac{1}{N_A^2}\frac{1}{N_B^2}\sum_{i,j,k,l} e^{i\boldsymbol{k}_1\cdot\boldsymbol{r}_i + \cdots + i\boldsymbol{k}_4\cdot\boldsymbol{r}_l} \\ &\times\langle (N_{A,i} - \langle N_{A,i}\rangle)(N_{A,j} - \langle N_{A,j}\rangle) \\ &\times (N_{B,k} - \langle N_{B,k}\rangle)(N_{B,l} - \langle N_{B,l}\rangle)\rangle. \end{aligned} \tag{31}$$

Thus, we find that in order to evaluate the trispectrum we are in turn required to evaluate the four-point cross-correlation function of counts in cells. Again, following Peebles (1980), we break this quadruple sum into five types of terms, each of which arises from a particular partitioning of the indices $(i, j, k, l)$. Full details are presented in the following sections, those not wishing to be embattled at this stage should skip ahead to Section 3.3.

#### *3.2.1 Terms* $(i \neq j \neq k \neq l)$

Terms in the sum with these indices correspond to contributions to the product from the connected four-point correlation function of the field. These terms can be rewritten as

$$\begin{aligned} \langle (N_{A,i} - \langle N_{A,i}\rangle)\cdots(N_{B,l} - \langle N_{B,l}\rangle)\rangle &= \bar{n}_A^2\bar{n}_B^2\delta V_i \cdots \delta V_l \\ &\times \left(\eta^{AABB}_{ijkl} + \xi^{AA}_{ij}\xi^{BB}_{kl} + \xi^{AB}_{ik}\xi^{AB}_{jl} + \xi^{AB}_{il}\xi^{AB}_{jk}\right), \end{aligned} \tag{32}$$

where for convenience we have introduced the abbreviated notation for the two-, three- and four-point correlation functions:

$$\xi_{ij} \equiv \xi(\boldsymbol{r}_i, \boldsymbol{r}_j);\ \zeta_{\rm ijk} \equiv \zeta(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k);\ \eta_{\rm ijkl} \equiv \eta(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k, \boldsymbol{r}_l).$$

On inserting the above expression into equation (31), transforming the sums over cells to volume integrals and using the statistical homogeneity of the correlation functions, we obtain the following expression:

$$\begin{aligned} {V_\mu}^2 \left\langle \delta^d_A(\boldsymbol{k}_1)\ldots\delta^d_B(\boldsymbol{k}_4)\right\rangle &= \frac{1}{V_\mu} T_{AABB}(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0} \\ &+ P_{AA}(\boldsymbol{k}_1)P_{BB}(\boldsymbol{k}_3)\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_2}\delta^K_{\boldsymbol{k}_3,-\boldsymbol{k}_4} \\ &+ P_{AB}(\boldsymbol{k}_1)P_{AB}(\boldsymbol{k}_2)\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_3}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_4} \\ &+ P_{AB}(\boldsymbol{k}_1)P_{AB}(\boldsymbol{k}_2)\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_4}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_3}, \end{aligned} \tag{33}$$

where we have used the short-hand notation $\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0} = \delta^K_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4,0}$, and the irreducible or connected trispectrum of the underlying continuous density field has been defined as $T(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_4) \equiv V_\mu{}^3\langle\delta(\boldsymbol{k}_1)\ldots\delta(\boldsymbol{k}_4)\rangle_c\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0}$. This obeys a Fourier relation with the irreducible four-point correlation function $\eta_{\rm ijkl}$.

### 3.2.2 *Terms* $(i \neq j \neq k\ = l) + perms$

There are six types of term that arise from the equivalence of two of the indices, and in order to evaluate these we are required to deal with products of the form

$$\begin{aligned}\left\langle (N_{A,i}-\langle N_{A,i}\rangle)(N_{A,j}-\langle N_{A,j}\rangle)(N_{B,k}-\langle N_{B,k}\rangle)^2\right\rangle \\ = \langle (N_{A,i}-\langle N_{A,i}\rangle)(N_{A,j}-\langle N_{A,j}\rangle)N_{B,k}\rangle, \end{aligned} \tag{34}$$

$$= \bar{n}_A^2\bar{n}_B\delta V_i\delta V_j\delta V_j\left(\xi^{AA}_{ij}+\zeta^{AAB}_{\rm ijk}\right), \tag{35}$$

where the second equivalence follows from the rules for the cross-moments in Section 2.3. Hence, on repeating this procedure for all possible ways of equivalencing two indices we arrive at six expressions. Then, on following a procedure similar to the evaluation of the cross-power spectrum, and introducing the bispectrum $B$ as

$$B(\boldsymbol{k}_1,\boldsymbol{k}_2)\delta^K_{\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3,0} \equiv V_\mu^2\,\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle\,, \tag{36}$$

and noting that $B$ and $\zeta$ are Fourier duals, we find that these terms can be written:

$$\begin{aligned} V_\mu{}^2\left\langle\delta^d_A(\boldsymbol{k}_1)\ldots\delta^d_B(\boldsymbol{k}_4)\right\rangle \\ = \left[\frac{1}{\bar{n}_B}P_{AA}(\boldsymbol{k}_1)+\frac{1}{\bar{n}_A}P_{BB}(\boldsymbol{k}_3)\right]\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_2}\delta^K_{\boldsymbol{k}_3,-\boldsymbol{k}_4} \\ + \left\{\frac{1}{\bar{n}_A}\left[P_{AB}(\boldsymbol{k}_1)+P_{AB}(\boldsymbol{k}_2)\right]\left[\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_4}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_3}+\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_3}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_4}\right]\right\} \\ + \left[\frac{1}{N_B}B_{AAB}(\boldsymbol{k}_1,\boldsymbol{k}_2)+\frac{1}{N_A}B_{ABB}(\boldsymbol{k}_3,\boldsymbol{k}_4)\right]\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0} \\ + \left\{\frac{1}{N_A}\left[B_{ABB}(\boldsymbol{k}_1,\boldsymbol{k}_3)+B_{ABB}(\boldsymbol{k}_1,\boldsymbol{k}_4)+B_{ABB}(\boldsymbol{k}_2,\boldsymbol{k}_3)\right.\right. \\ \left.\left. + B_{ABB}(\boldsymbol{k}_2,\boldsymbol{k}_4)\right]\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0}\right\}, \end{aligned} \tag{37}$$

where $B_{ABB}$ and $B_{AAB}$ are the cross-bispectra of the fields $A$ and $B$.

### 3.2.3 *Terms* $(i\ = j \neq k\ = l) + perms$

There are three terms of this form that arise in the quadruple sum, and theses involve evaluation of quantities of the form

$$\begin{aligned}\left\langle (N_{A,i}-\langle N_{A,i}\rangle)^2(N_{B,k}-\langle N_{B,k}\rangle)^2\right\rangle = \langle N_{A,i}N_{B,k}\rangle \\ = \bar{n}_A\bar{n}_B\delta V_i\delta V_k\left(1+\xi^{AB}_{ik}\right), \end{aligned} \tag{38}$$

where again we have used the relations for the cross-moments from Section 2.3. On repeating this procedure for the other two terms, and repeating the analysis as before, we find that these types of terms can be written together as

$$\begin{aligned} V_\mu{}^2\left\langle\delta^d_A(\boldsymbol{k}_1)\ldots\delta^d_B(\boldsymbol{k}_4)\right\rangle = \frac{\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_2}\delta^K_{\boldsymbol{k}_3,-\boldsymbol{k}_4}}{\bar{n}_A\bar{n}_B} \\ + \left[\frac{1}{\bar{n}_A^2}\left(\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_3}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_4}+\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_4}\delta^K_{\boldsymbol{k}_2,-\boldsymbol{k}_3}\right)\right] \\ + \frac{1}{\bar{n}_A\bar{n}_B V_\mu}P_{AB}(\boldsymbol{k}_1+\boldsymbol{k}_2)\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0} \\ + \left\{\frac{1}{\bar{n}_A^2 V_\mu}\left[P_{BB}(\boldsymbol{k}_1+\boldsymbol{k}_3)+P_{BB}(\boldsymbol{k}_1+\boldsymbol{k}_4)\right]\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0}\right\}. \end{aligned} \tag{39}$$

### 3.2.4 *Terms* $(i\ = j\ = k \neq l) + perms$

There are four possible types of term that arise from this combination of indices and each of these requires us to evaluate a product like

$$\begin{aligned}\left\langle (N_{A,i}-\langle N_{A,i}\rangle)^2(N_{B,i}-\langle N_{B,i}\rangle)(N_{B,l}-\langle N_{B,l}\rangle)\right\rangle \\ = \langle N_{B,i}(N_{B,l}-\langle N_{B,l}\rangle)\rangle \\ = \bar{n}_B^2\xi^{BB}_{il}\delta V_i\delta V_l. \end{aligned} \tag{40}$$

Hence, on repeating this for the four possible arrangements of the indices, and using the methods described for the previous terms, we find that all of these terms reduce to the following expression:

$$\begin{aligned} V_\mu{}^2\left\langle\delta^d_A(\boldsymbol{k}_1)\ldots\delta^d_B(\boldsymbol{k}_4)\right\rangle = \left\{\frac{1}{\bar{n}_A\bar{n}_B V_\mu}\left[P_{AB}(\boldsymbol{k}_1)+P_{AB}(\boldsymbol{k}_2)\right]\right. \\ \left. + \frac{1}{\bar{n}_A^2 V_\mu}\left[P_{BB}(\boldsymbol{k}_3)+P_{BB}(\boldsymbol{k}_4)\right]\right\}\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0}. \end{aligned} \tag{41}$$

### 3.2.5 *Terms* $(i\ = j\ = k\ = l)$

There is only one form for this type of term in the quadruple sum, and to evaluate it we are required to compute the quantity,

$$\left\langle (N_{A,i}-\langle N_{A,i}\rangle)^2(N_{B,i}-\langle N_{B,i}\rangle)^2\right\rangle = \langle N_{B,i}\rangle = \bar{n}_B\delta V_i.$$

Hence, this has the form

$$V_\mu{}^2\left\langle\delta^d_A(\boldsymbol{k}_1)\ldots\delta^d_B(\boldsymbol{k}_4)\right\rangle = \left\{\frac{1}{\bar{n}_A^2\bar{n}_B V_\mu}\right\}\delta^K_{\boldsymbol{k}_1+\cdots+\boldsymbol{k}_4,0}. \tag{42}$$

## 3.3 Expressions for the cross-power covariance

The summation of equations (33, 37, 39, 41, 42) gives the complete description of all the terms entering the cross-trispectrum of Fourier modes for samples $A$ and $B$. We may now use this to obtain the full non-Gaussian covariance of the cross-power spectrum in two different modes $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. To do this, we simply take equations (33, 37, 39, 41, 42), and set the arguments of the wave modes to be

$$(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3,\boldsymbol{k}_4) \rightarrow (\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_1,-\boldsymbol{k}_2).$$

This gives us the quantity $\langle\delta^d_A(\boldsymbol{k}_1)\delta^d_B(-\boldsymbol{k}_1)\delta^d_A(\boldsymbol{k}_2)\delta^d_B(-\boldsymbol{k}_2)\rangle$. Hence, the covariance is given by

$$\begin{aligned} C^d_{AB} = \frac{1}{V_\mu}T_{AABB}(\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_1,-\boldsymbol{k}_2) \\ + \left[P_{AA}(\boldsymbol{k}_1)+\frac{1}{\bar{n}_A}\right]\left[P_{BB}(\boldsymbol{k}_2)+\frac{1}{\bar{n}_B}\right]\delta^K_{\boldsymbol{k}_1,-\boldsymbol{k}_2} \end{aligned}$$

$$+\left[P_{AB}(\boldsymbol{k}_1)+\left\{\frac{1}{\bar{n}_A}\right\}\right]\left[P_{AB}(\boldsymbol{k}_2)+\left\{\frac{1}{\bar{n}_A}\right\}\right]\delta^K_{\boldsymbol{k}_1,\boldsymbol{k}_2}$$

$$+\frac{B_{AAB}(\boldsymbol{k}_1,\boldsymbol{k}_2)}{N_B}+\frac{1}{N_A}[B_{ABB}(-\boldsymbol{k}_1,-\boldsymbol{k}_2)+\{B_{ABB}(\boldsymbol{k}_1,-\boldsymbol{k}_2)$$

$$+B_{ABB}(\boldsymbol{k}_2,-\boldsymbol{k}_2)+B_{ABB}(\boldsymbol{k}_2,-\boldsymbol{k}_1)+B_{ABB}(\boldsymbol{k}_1,-\boldsymbol{k}_1)\}]$$

$$+\frac{P_{AB}(\boldsymbol{k}_1+\boldsymbol{k}_2)}{\bar{n}_A\bar{n}_B V_\mu}+\left\{\frac{1}{\bar{n}_A^2 V_\mu}[P_{BB}(\boldsymbol{0})+P_{BB}(\boldsymbol{k}_1-\boldsymbol{k}_2)]\right\}$$

$$+\left\{\frac{1}{\bar{n}_A^2 V_\mu}[P_{BB}(-\boldsymbol{k}_1)+P_{BB}(-\boldsymbol{k}_2)]\right\}$$

$$+\left\{\frac{1}{\bar{n}_A\bar{n}_B V_\mu}[P_{AB}(\boldsymbol{k}_1)+P_{AB}(\boldsymbol{k}_2)]\right\}+\left\{\frac{1}{\bar{n}_A^2\bar{n}_B V_\mu}\right\}. \quad (43)$$

Again, we remind the reader that the terms in curly brackets vanish for the case where samples $A$ and $B$ have no overlapping. It should also be noted that when samples $A$ and $B$ are equivalent, we recover the expressions for the covariance of the autopower spectrum (Meiksin & White 1999; Scoccimarro et al. 1999).

### 3.4 Band-power average covariance

The above formula gives us the full expression for the covariance in the cross-power spectrum per Fourier mode. In practice, the power is estimated by averaging over all wave modes in thin spherical shells in $k$-space – band powers. The band-power average power spectrum can be written:

$$\overline{P}_{AB}(k_i)=\frac{V_\mu}{V_{s,i}}\int_{V_{s,i}}\mathrm{d}^3k\ \langle\delta_A(\boldsymbol{k})\delta_B(-\boldsymbol{k})\rangle\,, \quad (44)$$

where the average is over the $k$-space shell $V_s$, of volume

$$V_{s,i}=\int_{k_i-\Delta k/2}^{k_i+\Delta k/2}\mathrm{d}^3k=4\pi k_i^2\Delta k\left[1+\frac{1}{12}\left(\frac{\Delta k}{k_i}\right)^2\right]. \quad (45)$$

The discretized form for the band power is

$$\overline{P}_{AB}(k)=\frac{V_\mu}{N_k}\sum_{j=1}^{N_k}\left\langle\delta_A(\boldsymbol{k}_j)\delta_B(-\boldsymbol{k}_j)\right\rangle, \quad (46)$$

where $N_k=V_s(k)/V_k$ is the total number of modes in the shell. $V_k=k_\mathrm{f}^3$ is the fundamental $k$-space cell volume and $k_\mathrm{f}=2\pi/L$ is the fundamental wave mode.

Likewise, the band-power averaged covariance can be written:

$$\overline{C}^d_{AB}[k_i,k_j]\equiv\frac{1}{V_{s,i}V_{s,j}}\int_{V_{s,i},V_{s,j}}\mathrm{d}^3k_1\mathrm{d}^3k_2\,C^d_{AB}[\boldsymbol{k}_1,\boldsymbol{k}_2]. \quad (47)$$

To obtain the full non-Gaussian band-power covariance, one then inserts equation (43) into the above expression, and this leads to

$$\begin{aligned}\overline{C}^d_{AB}[k_i,k_j]&=\frac{1}{V_\mu}\overline{T}_{AABB}[k_i,k_j]\\&+\frac{1}{N_k}\left[\left(\overline{P}_{AA}(k_i)+\frac{1}{\bar{n}_A}\right)\left(\overline{P}_{BB}(k_j)+\frac{1}{\bar{n}_B}\right)\right.\\&\left.+\left(\overline{P}_{AB}(k_i)+\left\{\frac{1}{\bar{n}_A}\right\}\right)\left(\overline{P}_{AB}(k_j)+\left\{\frac{1}{\bar{n}_A}\right\}\right)\right]\delta^K_{k_i,k_j}\\&+\frac{\overline{B}_{AAB}(k_i,k_j)}{N_B}+\frac{\overline{B}_{ABB}(k_i,k_j)}{N_A}+\left\{\frac{2}{N_A}\overline{B}_{ABB}(k_i,k_j)\right\}\\&+\frac{\overline{P}_{AB}[k_i,k_j]}{\bar{n}_A\bar{n}_B V_\mu}+\left\{\frac{1}{\bar{n}_A^2 V_\mu}\left[\overline{P}_{BB}[k_i,k_j]\right]\right\}\\&+\left\{\frac{1}{\bar{n}_A\bar{n}_B V_\mu}\left[\overline{P}_{AB}(k_i)+\overline{P}_{AB}(k_j)\right]\right\}\\&+\left\{\frac{1}{\bar{n}_A^2 V_\mu}\left[\overline{P}_{BB}(k_i)+\overline{P}_{BB}(k_j)\right]\right\}+\left\{\frac{1}{\bar{n}_A^2\bar{n}_B V_\mu}\right\}, \quad (48)\end{aligned}$$

where the bin-averaged trispectrum and bispectrum are

$$\overline{T}[k_i,k_j]\equiv\int_{V_{s,i},V_{s,j}}\frac{\mathrm{d}^3k_1}{V_{s,i}}\frac{\mathrm{d}^3k_2}{V_{s,j}}T(\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_1,-\boldsymbol{k}_2), \quad (49)$$

$$\overline{B}[k_i,k_j]\equiv\int_{V_{s,i},V_{s,j}}\frac{\mathrm{d}^3k_1}{V_{s,i}}\frac{\mathrm{d}^3k_2}{V_{s,j}}B(\boldsymbol{k}_1,\boldsymbol{k}_2,-\boldsymbol{k}_1-\boldsymbol{k}_2), \quad (50)$$

and where we introduced the function

$$\overline{P}[k_i,k_j]\equiv\int_{V_{s,i},V_{s,j}}\frac{\mathrm{d}^3k_1}{V_{s,i}}\frac{\mathrm{d}^3k_2}{V_{s,j}}P(|\boldsymbol{k}_1-\boldsymbol{k}_2|), \quad (51)$$

$$=\int_{V_{s,i},V_{s,j}}\frac{\mathrm{d}^3k_1}{V_{s,i}}\frac{\mathrm{d}^3k_2}{V_{s,j}}P(|\boldsymbol{k}_1+\boldsymbol{k}_2|). \quad (52)$$

We may now consider a number of interesting limiting cases of the above expressions. First, in the very large-scale limit $(k_i,k_j)\to 0$, and for CDM models that evolve from Gaussian random initial fluctuations, we have $\overline{P}\propto k\to 0$, and from gravitational instability (Bernardeau et al. 2002) we have $\overline{B}\propto\overline{P}^2\to 0$ and $\overline{T}\propto\overline{P}^3\to 0$, and the covariance becomes

$$\overline{C}^d_A[k_i,k_j]\approx\frac{1}{N_{k_i}}\frac{1}{\bar{n}_A\bar{n}_B}\delta^K_{k_i,k_j}+\left\{\frac{1}{\bar{n}_A^2\bar{n}_B V_\mu}\right\}. \quad (53)$$

Secondly, in the small-scale limit $(k_i,k_j)\gg k_\mathrm{f}=2\pi/L$, we have $N_k\gg 1$ and $\overline{P}\propto k^{-3}\to 0$, and again from gravitational instability we have $\overline{B}\propto\overline{P}^2\to 0$ and $\overline{T}\propto\overline{P}^3\to 0$, hence

$$\overline{C}^d_{AB}[k_i,k_j]\approx\left\{\frac{1}{\bar{n}_A^2\bar{n}_B V_\mu}\right\}. \quad (54)$$

The correlation matrix $\mathbf{C}$ is defined as the covariance matrix normalized by its diagonal components, i.e.

$$\mathbf{C}^d_{AB}[k_i,k_j]=\frac{\overline{C}^d_{AB}[k_i,k_j]}{\sqrt{\overline{C}^d_{AB}[k_i,k_i]\overline{C}^d_{AB}[k_j,k_j]}}, \quad (55)$$

and $\mathbf{C}[k_i,k_i]=1$ and $-1\le\mathbf{C}[k_i,k_j]\le 1$. Thus, for $(i\ne j)$ and in the large-scale limit we find

$$\begin{aligned}\mathbf{C}^d_{AB}[k_i,k_j]&\approx\left[\frac{(\bar{n}_A V_\mu)^2}{N_{k_i}N_{k_j}}+\bar{n}_A V_\mu\left(\frac{N_{k_i}+N_{k_j}}{N_{k_i}N_{k_j}}\right)+1\right]^{-1/2}.\\&\approx\frac{\sqrt{N_{k_i}N_{k_j}}}{\bar{n}_A V_\mu}\ll 1, \quad (56)\end{aligned}$$

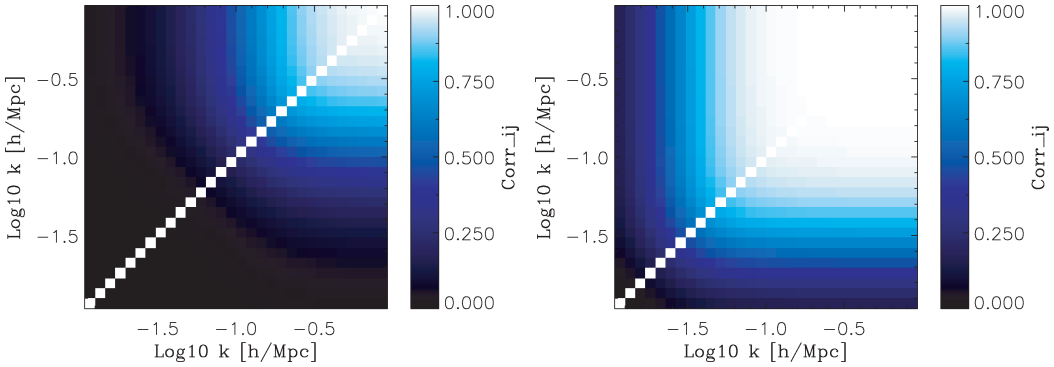


**Figure 1.** Theoretical predictions for the halo–halo autopower spectrum correlation matrix as a function of wavenumbers $k_i$ and $k_j$. Here, all of the covariance is generated by the discreteness terms and all non-Gaussian terms generated through gravitational instability have been set to zero, i.e. $T_c = B_c = 0$. Left-hand panel: dark matter haloes with masses $M \in [1.0 \times 10^{13}, 2.0 \times 10^{13}]\,h^{-1}\,\mathrm{M}_\odot$. Right-hand panel: dark matter haloes with masses $M \in [1.0 \times 10^{15}, 2.0 \times 10^{15}]\,h^{-1}\,\mathrm{M}_\odot$.

where the second equality obtains from assuming $\bar{n}_A V_\mu \to \infty$. Conversely, in the small-scale limit we find

$$\mathbf{C}^d_{AB}[k_i, k_j] \approx 1. \tag{57}$$

These last two expressions are important results. The first informs us that if one computes the autopower spectrum of a discrete sampling of the density field then for a standard CDM power spectrum the covariance matrix is diagonal on large scales provided $\bar{n} V_\mu \gg 1$. However, on small scales all of the Fourier modes inevitably become perfectly correlated, and this is completely independent of any non-Gaussian terms generated through gravitational instability. On the other hand, if there is no overlapping between samples $A$ and $B$ then there will be no off-diagonal covariance, since equation (54) vanishes.

We may demonstrate these statements more clearly by taking the Gaussian limit of equation (48),

$$\begin{aligned}\overline{C}^d_{AB}[k_i, k_j] = \frac{1}{N_k}\Bigg[&\left(\overline{P}_{AA}(k_i) + \frac{1}{\bar{n}_A}\right)\left(\overline{P}_{BB}(k_j) + \frac{1}{\bar{n}_B}\right) \\ &+ \left(\overline{P}_{AB}(k_i) + \left\{\frac{1}{\bar{n}_A}\right\}\right)\left(\overline{P}_{AB}(k_j) + \left\{\frac{1}{\bar{n}_A}\right\}\right)\Bigg]\delta^K_{k_i,k_j} \\ &+ \frac{\overline{P}_{AB}[k_i, k_j]}{\bar{n}_A \bar{n}_B V_\mu} + \left\{\frac{1}{\bar{n}_A^2 V_\mu}\left[\overline{P}_{BB}[k_i, k_j]\right]\right\} \\ &+ \left\{\frac{1}{\bar{n}_A^2 V_\mu}\left[\overline{P}_{BB}(k_i) + \overline{P}_{BB}(k_j)\right]\right\} \\ &+ \left\{\frac{1}{\bar{n}_A \bar{n}_B V_\mu}\left[\overline{P}_{AB}(k_i) + \overline{P}_{AB}(k_j)\right]\right\} + \left\{\frac{1}{\bar{n}_A^2 \bar{n}_B V_\mu}\right\}. \end{aligned} \tag{58}$$

For the case where samples $A$ and $B$ are identical, the above expressions reduce to

$$\begin{aligned}\overline{C}^d[k_i, k_j] = &\frac{2}{N_{k_i}}\left[\overline{P}(k_i) + \frac{1}{\bar{n}}\right]^2 \delta^K_{k_i,k_j} \\ &+ \frac{2}{\bar{n}^2 V_\mu}\left[\overline{P}[k_i, k_j] + \overline{P}(k_i) + \overline{P}(k_j)\right] + \frac{1}{\bar{n}^3 V_\mu}. \end{aligned} \tag{59}$$

Finally, since it will be of use later, we may also take the limit $\bar{n}_A V_\mu \to \infty$, giving

$$\begin{aligned}\overline{C}^d_{AB}[k_i, k_j] = \frac{1}{N_k}\Bigg[&\left(\overline{P}_{AA}(k_i) + \frac{1}{\bar{n}_A}\right)\left(\overline{P}_{BB}(k_j) + \frac{1}{\bar{n}_B}\right) \\ &+ \left(\overline{P}_{AB}(k_i) + \left\{\frac{1}{\bar{n}_A}\right\}\right)\left(\overline{P}_{AB}(k_j) + \left\{\frac{1}{\bar{n}_A}\right\}\right)\Bigg]\delta^K_{k_i,k_j}. \end{aligned} \tag{60}$$

and for $A$ equivalent to $B$

$$\overline{C}^d[k_i, k_j] = \frac{2}{N_{k_i}}\left(\overline{P}(k_i) + \frac{1}{\bar{n}}\right)^2 \delta^K_{k_i,k_j}. \tag{61}$$

Fig. 1 presents the correlation matrix for the halo–halo autopower spectrum generated using equation (59). In the left- and right-hand panels, we show the results for haloes with masses in the range $M \in [1.0, 2.0] \times 10^{13}\,h^{-1}\,\mathrm{M}_\odot$ and $M \in [1.0, 2.0] \times 10^{15}\,h^{-1}\,\mathrm{M}_\odot$, respectively. We evaluate the average bias and halo number density within each mass bin using the Sheth & Tormen (1999) models, and we find $\bar{n} = [1.87 \times 10^{-4}, 1.12 \times 10^{-7}]\,h^3\,\mathrm{Mpc}^{-3}$ and $b = [1.30, 5.85]$ for the two bins, respectively, and take $V_\mu = (1500)^3\,h^{-3}\,\mathrm{Mpc}^3$. In both cases, the matrix becomes fully correlated and the rare sample shows a much stronger correlation on larger scales than the higher abundance lower mass halo sample. On the other hand, if we were to compute the correlation matrix for the cross-power spectrum of the two halo samples then we would predict that the correlation matrix would be equivalent to the identity matrix.

Before we leave this section, it is interesting to note that, in the pure Gaussian limit, i.e. $\bar{n}_T P_T \gg 1$, the fractional variance in the cross-power is not simply dependent upon the number of available modes, but also the cross-correlation coefficient: $r_{AB}(k) \equiv P_{AB}(k_i)/\sqrt{P_{AA}(k_i)P_{BB}(k_i)}$. This can be seen directly from equation (58),

$$\left(\frac{\sigma_{\overline{P}_{AB}}}{\overline{P}_{AB}}\right)^2 = \frac{1}{N_k}\left(\frac{1}{r^2_{AB}} + 1\right). \tag{62}$$

The corresponding expression for the autopower spectrum is $\overline{\sigma}_P/\overline{P} = \sqrt{2/N_k} \propto k^{-1} V_\mu^{-1/2}$. However, when $r_{AB} = 1$ there is no difference and the fractional error scales with the survey volume in the usual way. In Section 6.2, we will show that for haloes and dark matter on the largest scales the cross-power approach offers only a modest improvement over the autopower method, implying that $r_{AB} \approx 1$.

## 4 EFFICIENCY OF ESTIMATORS

### 4.1 Comparing estimators

One might ask the following question: when should one apply the cross-power spectrum approach, instead of the usual autopower spectrum approach? In this section, we will attempt to answer this question. The main advantages of the cross-power approach are

most apparent when one probes the dependence of a given sample of tracers as a function of some physical parameter, i.e. the luminosity dependence of clustering or the mass dependence of the clustering of clusters. This statement can be more directly quantified if we consider the concept of estimator efficiency.

If we have two unbiased estimators $\mathcal{E}_1$ and $\mathcal{E}_2$ then the most efficient estimator of the two is said to be the one with the smallest variance: i.e. if $\mathrm{Var}[\mathcal{E}_1] < \mathrm{Var}[\mathcal{E}_2]$ then $\mathcal{E}_1$ will be considered to be a more efficient estimator than $\mathcal{E}_2$ (Barlow 1989). We need to modify this concept slightly since in comparing the cross- and autopower spectra we are not estimating the same thing, owing to the clustering bias. Instead, we will define the effective efficiency of the estimator through the signal-to-noise ratio (S/N): i.e. $\mathcal{E}_1$ will be considered to be a more efficient estimator than $\mathcal{E}_2$ if $\mathcal{E}_1/\sqrt{\mathrm{Var}[\mathcal{E}_1]} > \mathcal{E}_2/\sqrt{\mathrm{Var}[\mathcal{E}_2]} >$. Or in other words the estimator with the largest S/N will be the most efficient estimator.

On taking the limit $\bar{n}_A V_\mu \to \infty$ for equations (58) and (59), the covariance matrices are diagonal and so we may write the S/N for the auto and cross-power spectra as

$$\frac{(\mathrm{S/N})^2_{jj}}{N_k} = \frac{1}{2}\frac{\gamma_j^2}{\left(1+\gamma_j^2\right)}, \tag{63}$$

$$\frac{(\mathrm{S/N})^2_{ij}}{N_k} = \left[\frac{\gamma_i\gamma_j r_{ij}^2}{(\gamma_i+1)\left(\gamma_j+1\right)+(\sqrt{\gamma_i\gamma_j}r_{ij}+\{\delta\})^2}\right], \tag{64}$$

where we have introduced the following quantities:

$$\gamma_i \equiv \bar{n}_i P_{ii}, \tag{65}$$

$$r_{ij}^2 = P_{ij}^2/P_{ii}P_{jj}, \tag{66}$$

$$\alpha^2 \equiv \bar{n}_j/\bar{n}_i, \tag{67}$$

where we have taken the index $i$ to denote the high-density sample $A$ and $j$ to denote the low-density sample $B$. Taking the ratio of the above expressions gives us a simple test for the relative efficiency of the estimators,

$$\frac{(\mathrm{S/N})^2_{ij}}{(\mathrm{S/N})^2_{jj}} = 2r_{ij}^2\frac{\gamma_i}{\gamma_j} \times \left[\frac{(1+\gamma_j)^2}{(1+\gamma_i)(1+\gamma_j)+(\sqrt{\gamma_i\gamma_j}r_{ij}+\{\alpha\})^2}\right], \tag{68}$$

and we see that the relative efficiency does not depend explicitly on the number of available modes, nor the survey volume.

To proceed further we must specify samples $i$ and $j$ in more detail. Let us consider the case where sample $i$ is obtained from a set of unbiased high-density objects and where sample $j$ is obtained from a set of highly biased but rare objects. For this situation, we have $\bar{n}_i \gg \bar{n}_j$. Hence, $\alpha \to 0$. Further, we will assume that $\gamma_i \gg \gamma_j$. Hence, equation (68) simplifies to

$$\frac{(\mathrm{S/N})^2_{ij}}{(\mathrm{S/N})^2_{jj}} \approx 2r_{ij}^2\left[\frac{\gamma_j^2+2\gamma_j+1}{\gamma_j^2\left(1+r_{AB}^2\right)+\gamma_j}\right]. \tag{69}$$

On assuming that $r_{AB} \approx 1$, we finally find that

$$\frac{(\mathrm{S/N})^2_{ij}}{(\mathrm{S/N})^2_{jj}} \approx \left[\frac{2\gamma_j^2+4\gamma_j+2}{2\gamma_j^2+\gamma_j}\right] > 1. \tag{70}$$

This means that for examining the clustering properties of rare samples of objects it is more efficient to cross-correlate them with a high-density sample, rather than to compute their autopower spectrum.

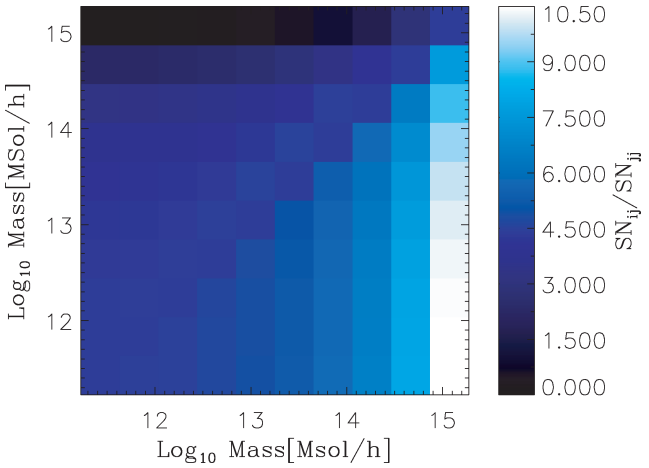


**Figure 2.** Relative signal-to-noise ratio matrix $(\mathrm{S/N})_{ij}/(\mathrm{S/N})_{jj}$ of the cross-power spectra of cluster samples in mass bin $i$ ($y$-axis) and mass bin $j$ ($x$-axis).

### 4.2 Example: improving estimates of cluster bias

Let us now provide a more concrete example. Consider a sample of dark matter clusters and suppose that we have both the redshift and an unbiased estimate of the cluster mass, i.e. through weak lensing, the Sunyaev–Zel'Dovich effect, etc., and that the clusters span the mass range $M \in [10^{11},\, 5\times10^{15}]\,h^{-1}\,\mathrm{M}_\odot$. We are interested in exploring the bias as a function of cluster mass, perhaps for use in constraining primordial non-Gaussianity as in Slosar et al. (2008). The sample may be subdivided into mass bins, and one may measure the autopower spectrum of each mass bin and also the cross-power spectra of the different mass bins.

Fig. 2 shows how the relative S/N as given by equation (68) varies as a function of the mass bins $i$ and $j$. Note that in the figure $i$ and $j$ represent the rows and columns of the matrix, respectively. When $i < j$, we find that there are significant advantages to be gained from computing the cross-power spectrum as opposed to the autopower spectrum especially for the case of high-mass haloes. For the case where $i > j$, naturally, the cross-power spectra are not optimal measures compared to the autospectrum. In generating this example, we used the dark matter halo mass function and bias formulation presented by Sheth & Tormen (1999), and we evaluate these for the same cosmological model as described in Table 1.

## 5 COVARIANCE OF THE CROSS-CORRELATION FUNCTION

As a corollary to our study of the cross-power spectrum, we extend our analysis to encompass the covariance of the cross-correlation function. We note that the autocorrelation covariance of dark matter and haloes on scales relevant for the Baryonic Acoustic Oscillations ($r \sim 100\,h^{-1}$ Mpc) was recently investigated in detail by Smith, Scoccimarro & Sheth (2008a) and Sánchez, Baugh & Angulo (2008). Here, we perform a similar study for the cross-correlation function.

In direct analogy with the analysis of power spectrum band powers, we may define the band-averaged cross-correlation function as

$$\bar{\xi}^{AB}(\boldsymbol{r}_i) = \frac{1}{V_s(r_i)}\int_{V_s(r_i)} \mathrm{d}^3r\,\xi^{AB}(r) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}P^{AB}(k)\bar{j}_0(kr_i), \tag{71}$$

**Table 1.** Parameters for the ZHORIZON simulations.

| Simulation | $\Omega_{\rm m}$ | $\Omega_{\rm w}$ | $\Omega_{\rm b}$ | $w_0$ | $\sigma_8$ | $n$ | $h$ | $N$ | $L(h^{-1}\,{\rm Mpc})$ | $m_{\rm p}(h^{-1}\,{\rm M}_\odot)$ | $N_{\rm sim}$ | $V_{\rm tot}\,(h^{-3}\,{\rm Gpc}^3)$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ZHORIZON | 0.25 | 0.75 | 0.04 | −1 | 0.8 | 1.0 | 0.7 | $750^3$ | 1500.0 | $5.55\times10^{11}$ | 30 | 101.25 |

*Note*. Columns are: density parameters for matter, dark energy and baryons; the equation of state parameter for the dark energy $P_{\rm w} = w\rho_{\rm w}$; normalization and primordial spectral index of the power spectrum; dimensionless Hubble parameter $H_0 = h100\,({\rm km\,s^{-1}\,Mpc^{-1}})$; number of particles, box size, particle mass, number of realizations and total simulation volume, respectively.

where $V_s$ is the radial shell of thickness $\Delta r$, over which the average is performed and this has volume

$$V_{s,i} = 4\pi r_i^2 \Delta r \left[1 + \frac{1}{12}\left(\frac{\Delta r}{r_i}\right)^2\right]. \tag{72}$$

For the second equality in equation (71), we have made use of the fact that $\xi \Leftrightarrow P$ are Fourier dual, and defined the zeroth-order bin-averaged spherical Bessel function as

$$\overline{j_0}(kr_i) \equiv \frac{r^2 j_1(kr)\big|_{r_1}^{r_2}}{r_i^2 k\Delta r\left[1+\frac{1}{12}\left(\frac{\Delta r}{r_i}\right)^2\right]}, \quad \begin{cases} r_2 = r_i + \Delta r/2 \\ r_1 = r_i - \Delta r/2 \end{cases}, \tag{73}$$

with $j_1(x) \equiv \sin x/x^2 - \cos x/x$ being the first-order spherical Bessel function. Similar to the bin-averaged covariance for the power (cf. equation 47), we may also define the bin-averaged cross-correlation covariance between bins $i$ and $j$:

$$\begin{aligned}\overline{C}^d_{\xi^{AB}} &\equiv {\rm Cov}\left[\overline{\xi}_i^{AB}, \overline{\xi}_j^{AB}\right] \\ &= \frac{1}{V_{s,i}V_{s,j}}\int_{V_{s,i},V_{s,j}} {\rm d}^3r_1{\rm d}^3r_2\, C^d_{\xi^{AB}},\end{aligned} \tag{74}$$

where $C^d_{\xi^{AB}} = {\rm Cov}[\xi^{AB}(\boldsymbol{r}_1), \xi^{AB}(\boldsymbol{r}_2)]$. On inserting our expression for the bin-averaged correlation function, we may rewrite the above expression as

$$\overline{C}^d_{\xi^{AB}} = \int \frac{{\rm d}^3k_1}{(2\pi)^3}\frac{{\rm d}^3k_2}{(2\pi)^3}\overline{j_0}(k_1r_i)\overline{j_0}(k_2r_j)C^d_{P_{\rm AB}}. \tag{75}$$

Thus, the cross-power covariance also gives us the cross-correlation covariance. It should also be noted that even if $C^d_{P_{AB}}$ is diagonal $\overline{C}^d_{\xi^{AB}}$ is not, since the spherical Bessel functions in the integrand effectively smooth the information across different scales.

The full non-Gaussian contributions to the correlation covariance can be calculated by substituting equation (43) into the above expression. On taking the continuum limit for the Kronecker delta symbols, i.e. $\delta^K_{\boldsymbol{k}_1,\boldsymbol{k}_2} \to \delta^D(\boldsymbol{k}_1-\boldsymbol{k}_2)(2\pi)^3/V_\mu$, rewriting the spherical Bessel functions as

$$j_0(kr) = \frac{1}{4\pi}\int d\Omega_{\boldsymbol{r}}\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}) \tag{76}$$

and using the Fourier relations between the $N$-point correlation functions and polyspectra, we then find that

$$\begin{aligned}\overline{C}^d_{\xi^{AB}} =& \int_{V_{s,i}V_{s,j}}\frac{{\rm d}^3r_1}{V_{s,i}}\frac{{\rm d}^3r_2}{V_{s,j}}\int\frac{{\rm d}^3y}{V_\mu}\eta_{AABB}(\boldsymbol{r}_1+\boldsymbol{y},\boldsymbol{r}_2,\boldsymbol{y}) \\ &+\int_{V_{s,i}V_{s,j}}\frac{{\rm d}^3r_1}{V_{s,i}}\frac{{\rm d}^3r_j}{V_{s,j}}\int\frac{{\rm d}^3y}{V_\mu}\left[\xi_{AA}(\boldsymbol{y})\xi_{BB}(\boldsymbol{r}_1+\boldsymbol{r}_2+\boldsymbol{y})\right. \\ &\left.+\xi_{AB}(\boldsymbol{y})\xi_{AB}(\boldsymbol{r}_1+\boldsymbol{r}_2+\boldsymbol{y})\right] \\ &+\frac{1}{V_\mu}\int_{V_{s,i}V_{s,j}}\frac{{\rm d}^3r_1}{V_{s,i}}\frac{{\rm d}^3r_j}{V_{s,j}}\left[\frac{1}{\bar{n}_B}\xi_{AA}(\boldsymbol{r}_1+\boldsymbol{r}_2)\right. \\ &\left.+\frac{1}{\bar{n}_A}\xi_{BB}(\boldsymbol{r}_1+\boldsymbol{r}_2)+\left\{\frac{2}{\bar{n}_A}\xi_{AB}(\boldsymbol{r}_1+\boldsymbol{r}_2)\right\}\right] \\ &+\left[\frac{1}{\bar{n}_A\bar{n}_B}+\left\{\frac{1}{\bar{n}_A^2}\right\}\right]\frac{\delta^K_{i,j}}{V_\mu V_s(i)}+\frac{\overline{\zeta}_{AAB}(r_i,r_j)}{N_B} \\ &+\frac{1}{N_A}\left[\overline{\zeta}_{ABB}(r_i,r_j)+\left\{2\overline{\zeta}_{ABB}(r_i,r_j)\right\}\right] \\ &+\frac{\overline{\xi}_{AB}(r_j)}{\bar{n}_A\bar{n}_BV_\mu V_{s,i}}\delta^K_{i,j}+\left\{\frac{\overline{\xi}_{BB}(r_i)}{\bar{n}_A^2V_\mu V_{s,i}}\delta^K_{i,j}\right\} \\ &+\left\{\frac{1}{\bar{n}_A\bar{n}_BV_\mu}\left[\frac{\overline{\xi}_{AB}(r_i)}{V_{s,j}}\delta^K_{j,1}+\frac{\overline{\xi}_{AB}(r_j)}{V_{s,i}}\delta^K_{i,1}\right]\right\} \\ &+\left\{\frac{1}{\bar{n}_A^2\bar{n}_BV_\mu}\frac{\delta^K_{i,1}}{V_{s,i}}\frac{\delta^K_{j,1}}{V_{s,j}}\right\}.\end{aligned} \tag{77}$$

Again, we may take the Gaussian ($\eta = \zeta = 0$) limit of the full expression and find

$$\begin{aligned}\overline{C}^d_{\xi^{AB}} =& \frac{1}{V_\mu}\int\frac{{\rm d}^3k}{(2\pi)^3}\overline{j_0}(kr_i)\overline{j_0}(kr_j) \\ &\times\left[\left(P_{AA}(k_1)+\frac{1}{\bar{n}_A}\right)\left(P_{BB}(k_1)+\frac{1}{\bar{n}_B}\right)\right. \\ &\left.+\left(P_{AB}(k_1)+\left\{\frac{1}{\bar{n}_A}\right\}\right)^2\right] \\ &+\frac{\overline{\xi}_{AB}(r_j)}{\bar{n}_A\bar{n}_BV_\mu V_{s,i}}\delta^K_{i,j}+\left\{\frac{\overline{\xi}_{BB}(r_i)}{\bar{n}_A^2V_\mu V_{s,i}}\delta^K_{i,j}\right\} \\ &+\left\{\frac{1}{\bar{n}_A\bar{n}_BV_\mu}\left[\frac{\overline{\xi}_{AB}(r_i)}{V_{s,j}}\delta^K_{j,1}+\frac{\overline{\xi}_{AB}(r_j)}{V_{s,i}}\delta^K_{i,1}\right]\right\} \\ &+\left\{\frac{1}{\bar{n}_A^2\bar{n}_BV_\mu}\frac{\delta^K_{i,1}}{V_{s,i}}\frac{\delta^K_{j,1}}{V_{s,j}}\right\}.\end{aligned} \tag{78}$$

The first term is the usual Gaussian plus Poisson expression and this leads to off-diagonal covariance through the spherical Bessel functions. The second and third terms contribute only to the diagonal variance; however, the fourth and fifth terms contribute to the off-diagonal variance along lines of zero lag and the last contributes only to the zero-lag term. Therefore in the Gaussian limit, whilst the

covariance matrix for the correlation functions is in general non-diagonal, the terms associated with the Poisson noise that lead to off-diagonal terms in the power spectrum covariance do not generate off-diagonal covariance in the correlation function. However, in the more general case we see that additional off-diagonal terms can be generated when we have non-zero $\eta$ and $\zeta$. Furthermore, for the case of the cross-correlation function of a non-overlapping samples, all of the terms in curly brackets vanish, and the covariance is significantly reduced.

Finally, on taking the limit $\bar{n}_A V_\mu \to \infty$ the covariance between band averages of the cross-correlation function reduces to

$$\overline{C}^d_{\xi_{AB}} \approx \frac{1}{V_\mu}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\overline{j_0}(kr_i)\overline{j_0}(kr_j)\Gamma(k) + \frac{\delta^K_{i,j}}{V_\mu V_s(r_i)}\left[\frac{1}{\bar{n}_A\bar{n}_B}+\left\{\frac{1}{\bar{n}^2_A}\right\}\right], \quad (79)$$

and where we introduced the useful function:

$$\Gamma(k) = P_{AA}(k_1)P_{BB}(k_1) + \left[\frac{P_{AA}(k_1)}{\bar{n}_B}+\frac{P_{BB}(k_1)}{\bar{n}_A}\right] + P^2_{AB}(k_1) + \left\{\frac{2P_{AB}(k_1)}{\bar{n}_A}\right\}. \quad (80)$$

Lastly, in the limit where $\bar{n}_A \equiv \bar{n}_B$, we recover the usual expression in the Gaussian limit (Smith et al. 2008a; Sánchez et al. 2008).

## 6 COMPARISON WITH *N*-BODY SIMULATIONS

In this section, we compare the counts-in-cells predictions for the covariance matrices of the mass–mass, halo–mass and halo–halo power spectra and correlation functions with results from a large ensemble of numerical simulations.

### 6.1 The ZHORIZON simulations

The Zürich Horizon, 'ZHORIZON', simulations are a large ensemble of pure CDM $N$-body simulations ($N_{\rm sim} = 30$, $N_{\rm part} = 750^3$), performed at the University of Zürich on the ZBOX2 and ZBOX3 supercomputers. The specific aim for these simulations is to provide high-precision measurements of cosmic structures on the scales of the order of $\sim 100\,h^{-1}$ Mpc and to also provide insight into the rarest fluctuations within the $\Lambda$ cold dark matter ($\Lambda$CDM) model that we should expect to find within the observable universe – the Horizon Volume.

Each numerical simulation was performed using the publicly available GADGET-2 code (Springel 2005), and followed the non-linear evolution under gravity of $N$ equal-mass particles in a co-moving cube of length $L$. All of the simulations were run within the same cosmological model, and the particular choice for the parameters was inspired by results from the *WMAP* experiment (Spergel et al. 2003, 2007; Komatsu et al. 2009) – the parameters for the simulations are listed in Table 1. The transfer function for the simulations was generated using the publicly available CMBFAST code (Seljak & Zaldarriaga 1996; Seljak et al. 2003), with high sampling of the spatial frequencies on large scales. Initial conditions were lain down at redshift $z = 50$ using the serial version of the publicly available 2LPT code (Scoccimarro 1998; Crocce, Pueblas & Scoccimarro 2006).

Dark matter halo catalogues were generated for all snapshots of each simulation using the friends-of-friends (FoF) algorithm (Davis et al. 1985), with the linking length parameter set to the standard $b = 0.2$; $b$ is the fraction of the interparticle spacing. For this, we used the fast parallel B-FOF code, kindly provided by V. Springel. The minimum number of particles for which an object was considered to be a bound halo was set to 30 particles. This gave a minimum host halo mass of $\sim 10^{13}\,\mathrm{M}_\odot\,h^{-1}$.

### 6.2 Results: band-power variances

Fig. 3 shows the results for the mean fractional error in the mass–mass (bottom panel), halo–mass (middle panel) and halo–halo (top panel) power spectra, as measured from the ZHORIZON simulations. The spectra were estimated for each simulation using the standard methods (Smith et al. 2003; Jing 2005; Smith, Sheth & Scoccimarro 2008b): particles and halo centres were interpolated on to a $1024^3$ cubical mesh, using the CIC algorithm (Hockney & Eastwood 1988);

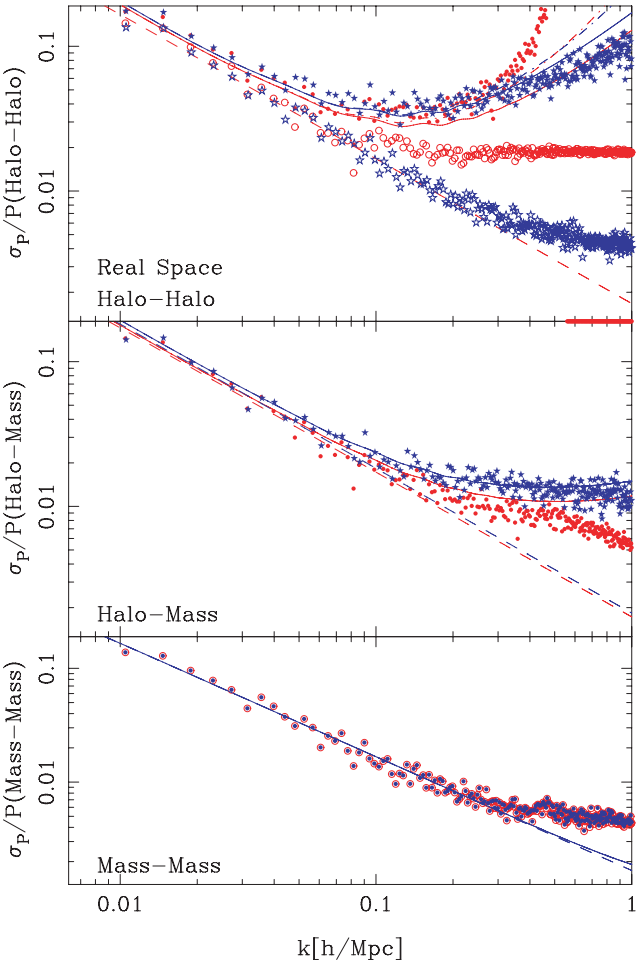


**Figure 3.** Comparison of the fractional variance in the halo and mass power spectra measured from the ZHORIZON simulations with theoretical predictions. The three panels show the square root of the bin-averaged diagonal elements of the covariance matrix, ratioed to the mean power in the bin as a function of the spatial frequency. From top to bottom, the panels show results for the halo–halo, halo–mass and mass–mass power spectra. In all panels, solid points denote results obtained after a standard shot-noise subtraction, and corresponding open points denote results prior to shot-noise subtraction. Dashed lines represent the pure Gaussian predictions. Solid lines denote theoretical predictions from the Gaussian plus standard Poisson noise theory. Dot–dashed lines denote the results from equations (58) and (59). In the top two panels, the (red) point symbols and (blue) star symbols denote haloes with masses in the range ($M > 1.0 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$) and ($1.0 \times 10^{13} < M < 2 \times 10^{13}\,h^{-1}\,\mathrm{M}_\odot$), respectively.

the Fast Fourier Transform of the discrete mesh was computed using the FFTW libraries (Johnson & Frigo 2008); the power in each Fourier mode was estimated and then corrected for the CIC charge assignment and these estimates were then bin averaged in spherical shells of thickness the fundamental frequency.

The halo–halo and halo–mass spectra were estimated for six bins in halo mass. The thickness of the mass bins was determined by estimating the S/N in each bin, and demanding that it should be in excess of 20. In the figure, we show the errors for an experiment of volume $\sim 3.4\,h^{-3}\,\mathrm{Gpc}^3$. For clarity, we only present results for the highest mass bin ($M > 10^{14}\,h^{-1}\,\mathrm{M}_\odot$, red point symbols) and for the lowest mass bin ($10^{13}\,h^{-1}\,\mathrm{M}_\odot < M < 1.38 \times 10^{13}\,h^{-1}\,\mathrm{M}_\odot$, blue star symbols) in our sample. The mean number densities in these bins are $\bar{n}_\mathrm{h} = \{2.42, 8.01\} \times 10^{-5}\,h^{-3}\,\mathrm{Mpc}^3$, respectively. The mass–mass and halo–halo power spectra were both corrected for shot noise by subtraction of $1/\bar{n} = V_\mu/N$ and $1/\bar{n}_\mathrm{h} = V_\mu/N_\mathrm{halo}$, respectively.

In the figure, the results for the shot-noise corrected and uncorrected spectra are represented as filled and empty symbols, respectively. The halo bias parameters were estimated from the cross-power and the shot-noise corrected autopower spectrum $\boldsymbol{b} = (b_\mathrm{NL}^{\mathrm{h}\delta}, b_\mathrm{NL}^{\mathrm{hh}})$ following the method in Smith et al. (2007). The measured values were found to be $\boldsymbol{b} = (2.803 \pm 0.015, 3.110 \pm 0.015)$ and $\boldsymbol{b} = (1.208 \pm 0.010, 1.479 \pm 0.011)$ for the highest and lowest mass bins, respectively. These estimates of the bias were used along with the ensemble average number densities in the mass bin to generate the theoretical predictions for the signal and its variance.

Considering Fig. 3 in more detail, we note that on the largest scales, $k \approx 0.01\,h\,\mathrm{Mpc}^{-1}$, the amplitudes of the fractional variances for all spectra are roughly equivalent. For the autospectra, this agreement is simply a consequence of the fact that when the signal is dominated by the sample variance the fractional errors in the spectra scale as $\sigma_P/P = (2/N_k)^{1/2} \propto V_\mu^{-1/2}$ (dashed lines in the figure). However, as we noted earlier, for the cross-spectrum, this near agreement also implies $r_{AB} \approx 1$.

For the matter power spectrum (bottom panel in Fig. 3), we see that this simple scaling appears to be preserved all the way to $k \approx 0.2\,h\,\mathrm{Mpc}^{-1}$, and here the errors are of the order of 1 per cent for this volume. The scaling at this point is broken and there is an excess of variance. This excess is not explained by the simple addition of the usual Poisson sampling error term (cf. equation 61), nor by the addition of the extra shot-noise terms from the full counts-in-cells covariance (cf. equation 59). However, in making these predictions, we have ignored all sources of variance generated through the non-linear gravitational mode coupling and it is likely that the excess error can be attributed to these (Meiksin & White 1999; Scoccimarro et al. 1999; Scoccimarro & Sheth 2002; Hamilton et al. 2006; Rimes & Hamilton 2006; Angulo et al. 2008a; Takahashi et al. 2009).

Considering the halo–mass cross-power spectra (middle panel in Fig. 3), we find that the scaling with the number of modes is broken on slightly larger scales than for matter ($k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$). At this point, the fractional error is of the order of $\sim$2 per cent. However, this time the increase in the error appears to be qualitatively described by equation (60) (solid line), although the error in the high-mass sample (red empty and filled circles) is slightly overpredicted. The additional source of variance in equation (58) does not change the predictions in any notable way. On smaller scales, ($k > 0.1\,h\,\mathrm{Mpc}^{-1}$), the fractional error drops to $\sim$1 per cent, and is only slightly larger than the error in the mass–mass spectrum. The excess theoretical error suggests that haloes and dark matter are not independent samples (as in sampling case i from Section 2.3), more, that haloes are some 'special' subsampling of the mass (similar to case ii), since we expect the Gaussian error to be an underestimate. This leads us to speculate that the halo–mass spectra also require a shot-noise correction.

Considering the halo–halo spectra (top panel of Fig. 3), we show results obtained with (solid symbols) and without (open symbols) the standard shot-noise subtraction. This clearly demonstrates the importance of this correction for this sample. In the case of the uncorrected spectra, it appears that the errors follow the scaling with the number of modes to high wavenumbers ($k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$), where the error is of the order of $\sim$2 per cent. In addition, we see that the standard theoretical predictions from equation (61) significantly overpredict the error, especially for the low-mass halo sample. However, after shot-noise subtraction, the sample variance scaling is actually broken on larger scales than for the cross-spectra, and the fractional error is of the order of $\sim$4–5 per cent. Somewhat surprisingly, these simple theoretical predictions provide a reasonable description of the variance and, as for the case of the matter–matter power spectrum, are an underestimate. If we now include the additional sources of variance from the full counts-in-cells covariance, as given by equation (59), then we now see that there is a significant increase in the errors for scales $k > 0.1\,h\,\mathrm{Mpc}^{-1}$. We have again neglected the gravitational model coupling variances, but it appears that most of the shape of this distribution is well captured by the non-Gaussianity of the sampling procedure. On comparison with Angulo et al. (2008a), we find a slight disagreement in that the Gaussian plus Poisson sampling model appears in reasonable agreement with the measurements.

Finally, we emphasize the fact that the fractional errors associated with the cross-power spectra are more than a factor of $\sim$2 times smaller than the corresponding errors for the halo autospectra on scales $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$. Thus, for experiments that wish to measure, for instance, galaxy bias as a function of luminosity, halo mass or galaxy type, one may gain a significant increase in S/N through use of the cross-correlation approach. The caveat being that the off-diagonal errors of the covariance matrix of the cross-power spectrum should be small. We will now explore this issue.

### 6.3 Results: mass–mass band-power correlation matrices

In Fig. 4, we present the correlation matrices for the mass–mass power spectrum as measured from the ZHORIZON simulations, where the correlation matrices are obtained from

$$\mathbf{C}[k_i, k_j] = \frac{C[k_i, k_j]}{\sqrt{C[k_i, k_i]C[k_j, k_j]}}. \tag{81}$$

For the correlation matrices, it was necessary to rebin the power spectra. This owed to the fact that when the power is averaged in shells of thickness the fundamental mode there are insufficient numbers of modes on large scales to produce a good S/N (Takahashi et al. 2009). We therefore chose to rebin the power by a factor of 4, and with the contribution from each $k$-shell being weighted by the number of modes in that shell. Lastly, we box car smoothed the matrices with a filter scale of width of 2 pixels.

In the left-hand panel of Fig. 4, we show the correlation matrices obtained from the power spectra without any shot-noise correction. It can clearly be seen that, going from large to small scales, there is a buildup of power correlations between neighbouring modes and for the smallest scales considered, the matrix appears perfectly correlated ($\mathbf{C} = 1$). In the right-hand panel, we show the same, but this time the matrix was generated from the shot-noise corrected power spectra. There are only small differences. It is likely that this result owes to the combination of two facts: first, the

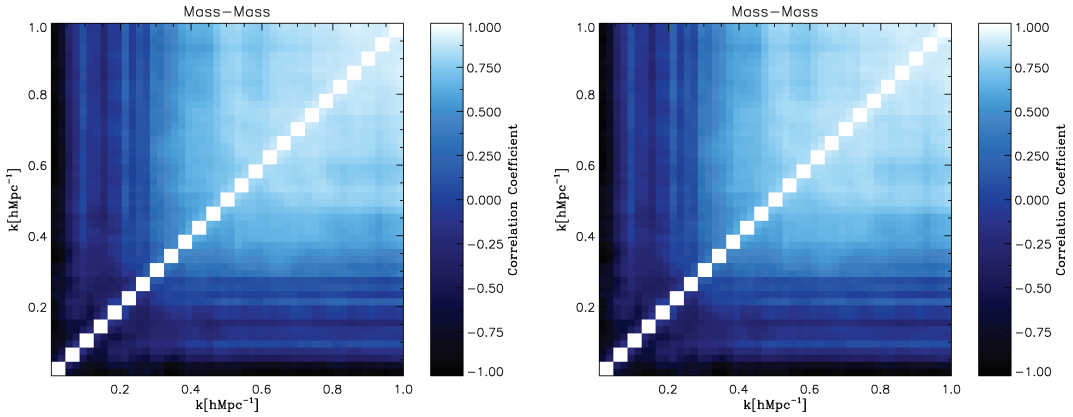


**Figure 4.** Mass–mass power correlation matrices measured from the ZHORIZON simulations. Left-hand panel: result before shot-noise correction. Right-hand panel: result after shot-noise correction.

number density of dark matter particles is sufficiently high to render the shot-noise contributions to the covariance of negligible importance (cf. Fig. 1); secondly, there is no variation in the number density of dark matter particles between realizations that might introduce additional variance (the importance of this will become clear in the next sections). Therefore, it is likely that the correlations are purely derived from the gravitational model coupling (for a recent and detailed study of the matter power spectrum covariance arising due to gravitational instability, see Takahashi et al. 2009).

### 6.4 Results: halo–halo band-power correlation matrices

Fig. 5 presents the results for the halo–halo autopower spectrum correlation matrices. The top two panels show the results obtained from the power spectra without shot-noise corrections. The left-hand panel shows results for cluster mass haloes ($M > 10^{14}\,h^{-1}\,\mathrm{M}_{\odot}$) and the right for group mass haloes ($10^{13} > M[\,h^{-1}\,\mathrm{M}_{\odot}] > 1.38 \times 10^{13}$). We see that the degree of correlation appears strongly dependent on both the halo mass range considered and also the scale considered, with the high-mass halo sample having significantly

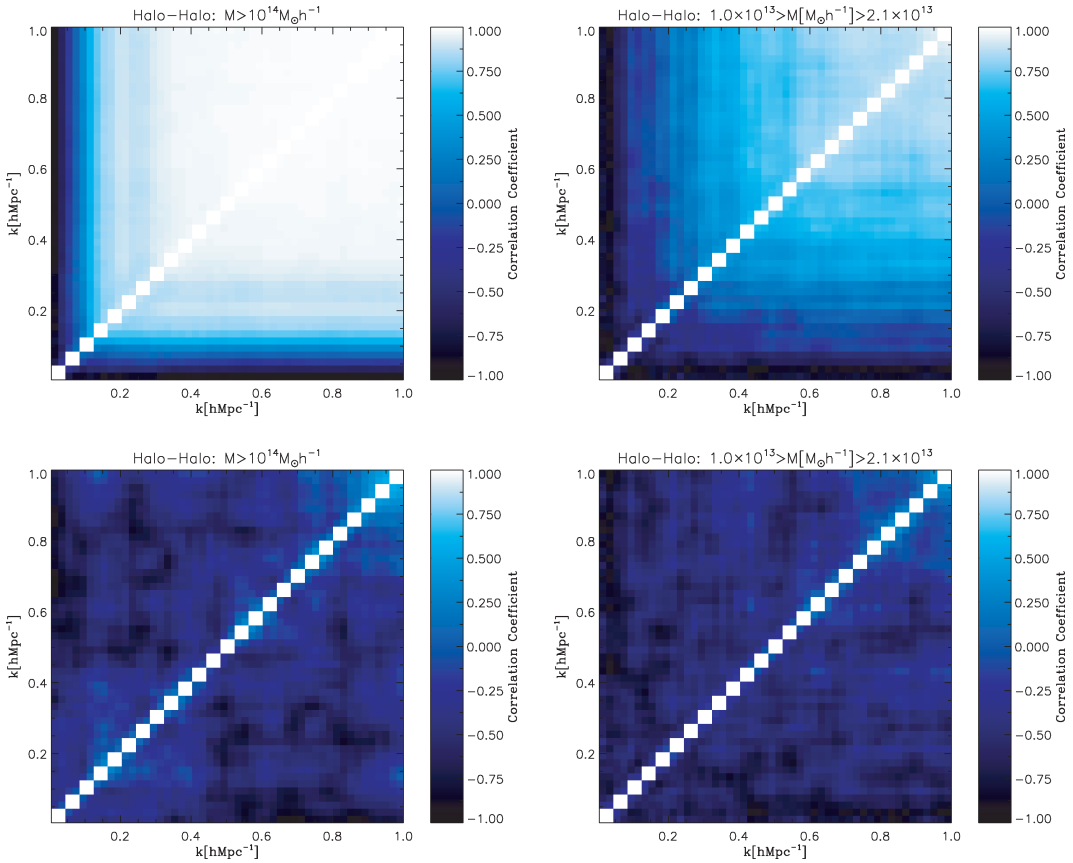


**Figure 5.** Halo–halo power spectrum correlation matrices measured from the ZHORIZON simulations. Top panels: results for power spectra without any correction for shot noise. Bottom panels: results after shot-noise correction. Left-hand column: results for the cluster mass halo sample ($M > 10^{14}\,h^{-1}\,\mathrm{M}_{\odot}$). Right-hand column: results for group mass haloes [$10^{13} > M(\,h^{-1}\,\mathrm{M}_{\odot}) > 1.38 \times 10^{13}$].

stronger off-diagonal correlations than for the lower mass sample for a given scale. Both matrices show significantly more correlation than was found for the dark matter.

In the bottom two panels of Fig. 5, we show the same matrices, but this time constructed from the shot-noise corrected power spectra. The difference is remarkable – the strong off-diagonal correlations that are present in the upper panels has been almost completely suppressed. The shot-noise corrected covariance matrix may be written in terms of the shot-noise uncorrected covariance as

$$\overline{C}^{c}_{\rm hh}[k_i, k_j] = \left\langle \left( \overline{P}^{d}_{\rm hh}(k_i) - \frac{1}{\bar{n}_{\rm h}} \right) \left( \overline{P}^{d}_{\rm hh}(k_j) - \frac{1}{\bar{n}_{\rm h}} \right) \right\rangle - \left\langle \overline{P}^{d}_{\rm hh}(k_i) - \frac{1}{\bar{n}_{\rm h}} \right\rangle \left\langle \overline{P}^{d}_{\rm hh}(k_j) - \frac{1}{\bar{n}_{\rm h}} \right\rangle, \quad (82)$$

$$= C^{d}_{\rm hh}[k_i, k_j] - {\rm Covar}\left[ \frac{1}{\bar{n}_{\rm h}}, \overline{P}^{d}_{\rm hh}(k_i) \right] - {\rm Covar}\left[ \frac{1}{\bar{n}_{\rm h}}, \overline{P}^{d}_{\rm hh}(k_j) \right] + {\rm Var}\left[ \frac{1}{\bar{n}_{\rm h}} \right]. \quad (83)$$

If the number density of the tracer sample does not vary between realizations then the shot corrected and uncorrected covariance matrices are identical. However, if it does then we see that there are additional sources of covariance that are introduced due to the coupling between the amplitude of the halo–halo power spectrum and the mass function of haloes, and from the variance of the number density. In order for the subtraction of shot-noise to result in a diagonal correlation matrix, it requires that the cross-correlation between the halo number counts and the halo power spectrum cancels with the off-diagonal contributions to $C^{d}_{\rm hh}[k_i, k_j]$. It is beyond the scope of this current work to illuminate this issue further and it shall remain as a topic for future investigation. One caveat to the above results is that it is well known that the standard shot-noise correction is too strong for haloes, since it results in negative power on small scales (Smith et al. 2007). It is therefore likely that this will have some impact on the covariance matrix.

Lastly, we now see that for the matter power spectra, since the number density of dark matter particles does not vary between realizations, we must have $\overline{C}^{c}_{\delta\delta}[k_i, k_j] = \overline{C}^{d}_{\delta\delta}[k_i, k_j]$.

### 6.5 Results: halo–mass band-power correlation matrices

In Fig. 6, we show the correlation matrices for the halo–mass cross-power spectra. The left-hand panel shows results for cluster mass haloes ($M > 10^{14}\,h^{-1}\,{\rm M}_{\odot}$) and the right for group mass haloes ($10^{13} > M[\,h^{-1}\,{\rm M}_{\odot}] > 1.38 \times 10^{13}$). Similar to the halo autopower correlation matrix, we see that the degree of correlation appears strongly dependent on both halo mass and scale. Interestingly, we note that whilst the spectra from the high-mass sample show more band-power correlation than for the dark matter the lower mass halo sample appears to show less. This further recommends the cross-spectra approach for further investigation as an improved estimator for large-scale structure.

All of the above matrices serve to warn us that, whilst the Gaussian plus Poisson model describes the diagonal errors reasonably well, it fails to capture the buildup of correlations between Fourier modes. To describe the above results, one must model both the full non-Gaussian trispectrum generated by gravitational mode-coupling (Scoccimarro et al. 1999; Takahashi et al. 2009) and, as we have shown in this paper, the covariance introduced by the point sampling for the mass tracers.

### 6.6 Results: band-correlation function variances

As a final study, we now consider the correlation function errors. The main advantage of the configuration space is that the constant shot-noise correction, which is necessary for the power spectra, is not required here. This follows from the fact that the Fourier transform of a constant gives a delta function at zero lag. However, as was described in Section 5, the shot-noise corrections do affect the correlation function errors.

In Fig. 7, we present measurements from the ensemble of ZHORIZON simulations for the fractional errors on the mass–mass (bottom panel), the halo–mass (middle panel) and the halo–halo (top panel) correlation functions. Again, we only show results for the highest and lowest bins in halo mass. The correlation functions were generated using the DUALTREETWOPOINT code, which is a parallel, tree-based algorithm and is described more fully in Smith et al. (in preparation). For the dark matter sample, we used roughly $\sim 4 \times 10^6$ particles, subsampled from the available $\sim 4 \times 10^8$ for each estimate.

The main result to note from this analysis is that, whilst for the power spectrum on large scales the fractional error is the same irrespective of tracer, this is not the case for the correlation function. We note that for $r > 20\,h^{-1}$ Mpc the halo–mass cross-correlation appears to be a more efficient estimator than the simple autocorrelation function, by almost a factor of $\sim 2$. To make this statement more concrete, we should include the off-diagonal errors in the

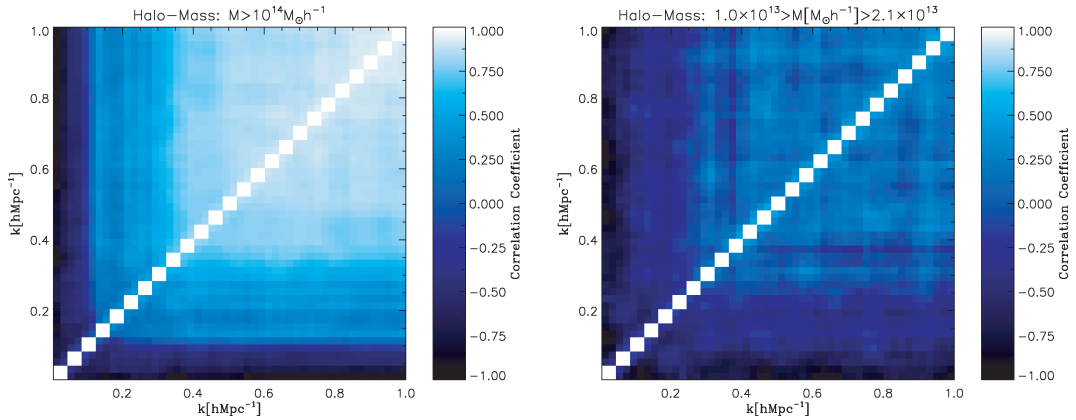


**Figure 6.** Halo–mass cross-power spectrum correlation matrices measured from the ZHORIZON simulations. Left-hand panel: results for a cluster mass halo sample ($M > 10^{14}\,h^{-1}\,{\rm M}_{\odot}$). Right-hand panel: results for a group mass halo sample [$10^{13} > M(\,h^{-1}\,{\rm M}_{\odot}) > 1.38 \times 10^{13}$].

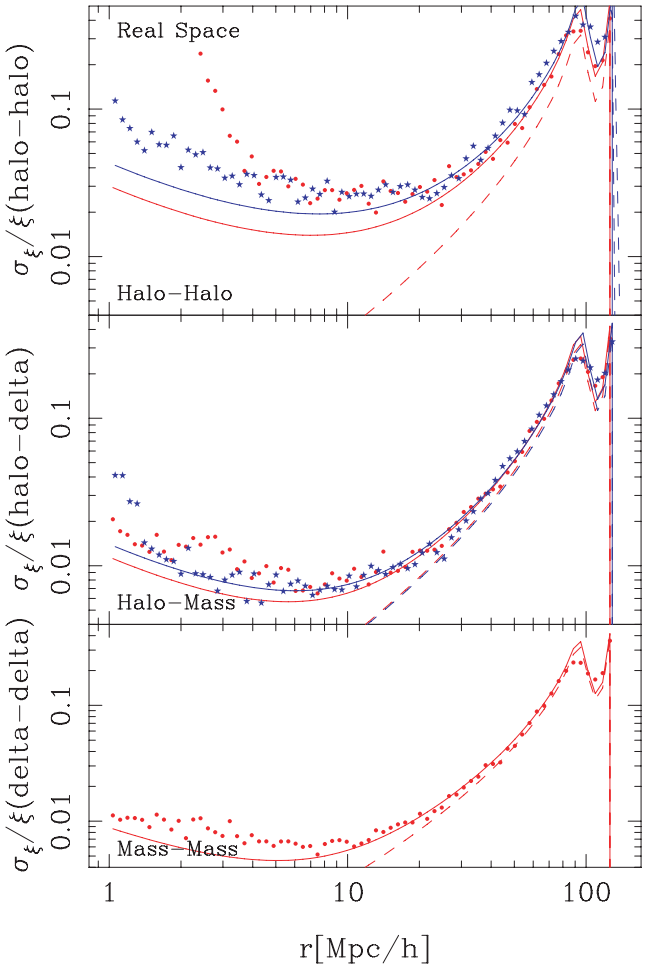


**Figure 7.** Comparison of the fractional variance of mass and halo correlation functions as measured from the ZHORIZON simulations with theoretical predictions. Similar to Fig. 3, the three panels show the standard deviation in the bin-averaged correlation functions, ratioed to the mean correlation function, as a function of the spatial scale. From top to bottom, the panels show results for the halo–halo, halo–mass and mass–mass correlations. Symbols show estimates measured from the $N$-body simulations. In the top two panels, the (red) point symbols and the (blue) star symbols denote haloes with masses in the range ($M > 1.0 \times 10^{14}\,h^{-1}\,\mathrm{M}_\odot$) and [$1.0 \times 10^{13} < M(h^{-1}\,\mathrm{M}_\odot) < 2 \times 10^{13}$], respectively. Again, the solid lines represent the theoretical predictions from the Gaussian plus Poisson sampling theory. Dashed lines represent the pure Gaussian predictions.

calculation of S/N. However, from our discussion in Section 5, we expect that the off-diagonal errors are also reduced. We will reserve this for a future work.

Another important point to note is that in nearly all cases the theoretical predictions for the Gaussian plus Poisson sampling error estimates are an underestimate of the measured errors, especially on scales ($r < 20\,h^{-1}$ Mpc). The predictions being worst for the autocorrelation function for the high-mass halo sample, and this is in agreement with the power spectrum results from the previous section.

The errors in the autocorrelation functions were previously investigated in numerical simulations by Smith et al. (2008a) and Sánchez et al. (2008), who showed that the Gaussian plus Poisson model provided a good description at the scale of the Baryonic Acoustic Oscillations ($r \sim 100\,h^{-1}$ Mpc). Our results extend this analysis to the cross-correlation functions. Also, the range of investigated scales is extended to smaller scales by more than one order of magnitude.

## 7 CONCLUSIONS

In this paper, we have performed a detailed investigation of the errors associated with auto and cross-power spectra and also the cross-correlation function of different tracers of the density field.

In Section 2, we developed the counts-in-cells approach for a multitracer approach to the clustering statistics. We showed that not all cross-power spectra are free from a shot-noise correction, and that the precise correction one should apply depends on the sampling distribution function.

In Section 3, we gave a derivation of the full non-Gaussian covariance matrix for the cross-power spectrum, including all the sources of variance that arise from the Poisson sampling of the mass tracers and this extends the standard results (Meiksin & White 1999; Scoccimarro et al. 1999; Cohn 2006; Hamilton et al. 2006). We showed that, for the case of Poisson sampling of Gaussian fluctuations there were terms that contributed to the off-diagonal terms of the covariance matrix. We showed that in the small-scale limit $k \to \infty$ these terms dominate over all other sources of variance (including Non-Gaussian terms generated from gravitational mode-coupling) and the covariance matrix becomes perfectly correlated.

In Section 4, we investigated the efficiency of the cross-power spectrum. We used the relative S/N of two different estimators as a diagnostic for efficiency. For the case where a high-density sample of tracers was cross-correlated with a low-density sample, it was shown that the former approach was a more efficient estimator than the case where one simply autocorrelates the low-density sample. As an example, we showed that for the determination of cluster bias, the cross-power spectrum approach would yield significant gains in S/N. Other uses are improving estimates of the luminosity dependence of galaxy bias.

In Section 5, we explored the covariance of auto and cross-correlation functions. It was shown that whilst the correlation function covariance matrix in general is not diagonal, the discreteness terms that led to off-diagonal covariance in the power spectrum do not generate off-diagonal elements in the correlation function covariance. Thus, the correlation function covariance matrix appears easier to understand and model than the power spectrum covariance.

In Section 6, we used a large ensemble of $N$-body simulations, to obtain estimates of the power spectrum and correlation function errors. We showed for the fractional errors on the mass–mass halo-mass and halo–halo spectra that the numerical results were in reasonably good agreement with the Gaussian plus Poisson sampling model, but the measurements showed larger variance than the theory. It was also shown that in the limit of large scales and in the case that Poisson error is not dominant the fractional errors for all spectra are equivalent, since they are simply $\propto k^{-1} V_\mu^{-1/2}$.

We investigated the correlation matrix for the mass–mass power spectrum, and confirmed that there were strong correlations between different band-powers (Meiksin & White 1999; Scoccimarro et al. 1999; Takahashi et al. 2009). We showed that correcting the spectra for shot-noise does not change the correlation matrix significantly. We investigated the halo–halo autopower covariance matrix without applying a correction for shot-noise. We showed that the degree of correlation increased with the mass of the halo sample considered and that the matrices showed more band-power correlation than for the dark matter. We then estimated the covariance from the shot-noise corrected spectra and found that the off-diagonal errors were dramatically reduced, almost decorrelating individual band

powers. We conjectured that this arises from the subtraction of the covariance between halo number density and the halo–halo power spectrum and also the variance in the halo number density. We investigated the cross-power correlation matrix for haloes and dark matter and showed that the correlations were reduced compared to the shot-noise uncorrected halo–halo matrices and for the lowest mass halo sample were less correlated than the dark matter. We investigated the errors in configuration space, and showed that there was a significant gain in S/N on all scales from using the cross-correlation function of haloes and dark matter as opposed to simply examining the autocorrelation function of haloes.

We conclude that, for certain cases, the cross-spectra and cross-correlation functions are more efficient probes for the large-scale structure, than the standard autospectra and autocorrelation function approaches that are widely in use. These cases concern studies aiming to measure: the luminosity dependence of the galaxy bias (Norberg et al. 2002; Tegmark et al. 2004a); the cluster bias as a function of mass and hence constrain the degree of primordial non-Gaussianity in the early Universe (Dalal et al. 2008; Slosar 2009; Pillepich et al. 2008; Desjacques et al. 2009).

## ACKNOWLEDGMENTS

RES acknowledges R. Angulo, V. Desjacques, C. Porciani, R. Scoccimarro, U. Seljak and R. Sheth for useful discussions, and A. Saintonge for help with IDL. RES kindly thanks L. Marian for comments on the draft. RES thanks V. Springel for making public his GADGET-2 code and for providing his B-FOF halo finder, R. Scoccimarro for making public his 2LPT initial conditions code and U. Seljak and M. Zaldarriaga for making public their CMBFAST code. RES acknowledges support from a Marie Curie Reintegration Grant and the Swiss National Foundation.

## REFERENCES

Angulo R. E., Baugh C. M., Frenk C. S., Lacey C. G., 2008a, MNRAS, 383, 755
Angulo R. E., Baugh C. M., Lacey C. G., 2008b, MNRAS, 387, 921
Barlow R., 1989, in Barlow R., ed., Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences. Wiley, New York, p. 71
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bernstein G. M., 1994, ApJ, 424, 569
Cohn J. D., 2006, New Astron., 11, 226
Cole S. et al., 2005, MNRAS, 362, 505
Coles P., 1993, MNRAS, 262, 1065
Cooray A., Hu W., 2001, ApJ, 554, 56
Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369
Dalal N., Doré O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Desjacques V., Seljak U., Iliev I. T., 2009, MNRAS, 396, 85
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447
Grossi M., Verde L., Carbone C., Dolag K., Branchini E., Iannuzzi F., Matarrese S., Moscardini L., 2009, MNRAS, 398, 321
Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188
Hockney R. W., Eastwood J. W., 1988, Computer Simulation Using Particles. Hilger, Bristol
Hütsi G., Lahav O., 2008, A&A, 492, 355
Jing Y. P., 2005, ApJ, 620, 559
Jing Y. P., Suto Y., Mo H. J., 2007, ApJ, 657, 664
Johnson S., Frigo M., 2008, FFTW Software v.3.2, http://www.fftw.org/
Komatsu E. et al., 2009, ApJS, 180, 330
McDonald P., Seljak U., 2008, preprint (arXiv:0810.0323)
Meiksin A., White M., 1999, MNRAS, 308, 1179
Norberg P. et al., 2002, MNRAS, 332, 827
Padmanabhan N., White M., Norberg P., Porciani C., 2008, MNRAS, submitted (arXiv:0802.2105)
Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Percival W. J. et al., 2001, MNRAS, 327, 1297
Percival W. J. et al., 2007, ApJ, 657, 51
Pillepich A., Porciani C., Hahn O., 2008, preprint (arXiv:0811.4176)
Rimes C. D., Hamilton A. J. S., 2006, MNRAS, 371, 1205
Sánchez A. G., Baugh C. M., Angulo R., 2008, MNRAS, 390, 1470
Scoccimarro R., 1998, MNRAS, 299, 1097
Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629
Scoccimarro R., Zaldarriaga M., Hui L., 1999, ApJ, 527, 1
Sefusatti E., Crocce M., Pueblas S., Scoccimarro R., 2006, Phys. Rev. D, 74, 023522
Seljak U., 2009, Phys. Rev. Lett., 102, 021302
Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437
Seljak U., Sugiyama N., White M., Zaldarriaga M., 2003, Phys. Rev. D, 68, 083507
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Slosar A., 2009, J. Cosmol. Astropart. Phys., 03, 004
Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, J. Cosmology Astropart. Phys., 8, 31
Smith R. E. et al., 2003, MNRAS, 341, 1311
Smith R. E., Scoccimarro R., Sheth R. K., 2007, Phys. Rev. D, 75, 063512
Smith R. E., Scoccimarro R., Sheth R. K., 2008a, Phys. Rev. D, 77, 043525
Smith R. E., Sheth R. K., Scoccimarro R., 2008b, Phys. Rev. D, 78, 023523
Spergel D. N. et al., 2003, ApJS, 148, 175
Spergel D. N. et al., 2007, ApJS, 170, 377
Springel V., 2005, MNRAS, 364, 1105
Stirling A. J., Peacock J. A., 1996, MNRAS, 283, L99
Takahashi R. et al., 2009, ApJ, 700, 479
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Tegmark M. et al., 2004a, ApJ, 606, 702
Tegmark M. et al., 2006, Phys. Rev. D, 74, 123507
Tegmark M. et al., 2004b, Phys. Rev. D, 69, 103501
Wake D. A., Croom S. M., Sadler E. M., Johnston H. M., 2008, MNRAS, 391, 1674
White M., Song Y.-S., Percival W. J., 2009, MNRAS, 397, 1348

This paper has been typeset from a TEX/LATEX file prepared by the author.